ABSTRACT

Title of Dissertation: WHAT MAKES A PATHOGEN?
GENETIC AND STRUCTURAL HETEROGENEITY
OF NEISSERIAL LIPOOLIGOSACCHARIDE

Ellen Therese O'Connor, Doctor of Philosophy, 2006

Dissertation directed by: Dr. Daniel C. Stein
Professor
Department of Cell Biology and Molecular Genetics

The genus *Neisseria* consists of both pathogenic and commensal species that colonize mucosal niches. Specific structures of neisserial lipooligosaccharide (LOS) expressed by the gonococcus and meningococcus have been shown to play a role in pathogenesis. Commensal *Neisseria* also produce LOS, but can express additional structures not found in the pathogenic strains. *N. sicca* 4320 differs from other commensal *Neisseria* because it caused disease in a seemingly healthy individual. Therefore, this strain was used to research the extent to which a non-pathogen differs from a pathogen. A bioinformatic examination of the nucleotide sequence of the 4320 chromosome revealed strong homologies with *N. meningitidis* MC58. These findings suggest that 4320 is a member of the *Neisseria* and that commensal and pathogenic *Neisseria* have similar genomic content. Through MALDI-TOF, exoglycosidase digestion, and MS^n analyses, 4320 was shown to express both LOS and LPS. The 4320 LOS resembles that of *Haemophilus sp.* in that it contained 3 heptoses. The LPS consists

of a *N*-acetylglucosamine and rhamnose disaccharide repeat that is not attached to a lipid A, a novel molecule.

PEA transfer onto pathogenic *Neisseria* LOS was examined, as strains expressing PEA at 3-HepII as opposed to 6-HepII are more serum resistant. Southern hybridization was used to identify the presence of phosphoethanolamine (PEA) transfer genes in *N. sicca* strains. 4320 showed differences in the presence of Lpt3 transfer genes compared to other strains within the species. Lpt3 was isolated and biochemically shown to mediate the addition of PEA to 3-HepII of LOS that was constructed to lack the 3-HepII decoration. *lpt3* nucleotide sequence analysis was done on a diverse group of gonococcal isolates, and phylogenetic analysis showed a high degree of sequence divergence. These data support the idea that the presence of PEA at 3-HepII of LOS increases the pathogenic ability of the *Neisseria*, as *lpt3* was only found in *N. sicca* strains isolated from fatal endocarditis cases.

WHAT MAKES A PATHOGEN?
GENETIC AND STRUCTURAL HETEROGENEITY
OF NEISSERIAL LIPOOLIGOSACCHARIDE

by

Ellen Therese O'Connor

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the degree requirements for the degree of
Doctor of Philosophy
2006

Advisory Committee:

Professor Daniel C. Stein, Chair/Advisor
Professor Philip DeShong
Associate Professor Jeffrey DeStefano
Associate Professor Richard Stewart
Associate Professor Wenxia Song

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Daniel Stein, for guiding and encouraging me for the past five years. I would also like to thank members of my dissertation committee including Dr. Spencer Benson, Dr. Norm Hansen, Dr. Philip DeShong, Dr. Jeffrey DeStefano, Dr. Rick Stewart, and Dr. Wenxia Song. I would not have made it this far without your interesting questions and ideas.

I also would like to thank Dr. Andrezj Piekarowicz. I enjoyed our talks about not only biochemistry, but about Polish culture. Thank you also to the mass spec collaborators, Dr. Vern Reinhold and Dr. Mac Griffiss. Without your knowledge, this project would not have yielded the results it did. Also, thanks needs to go out to all of my CBMG family in no specific order Lili, Jamie, Nate, Mike, Annie C., Meredith, Sam, Sean, Chris, Julie, Esteban, Hwalih, Mark, Adriana, Elizabeth, Jackie, Michelle, Kathryn, Zenab, John, Larry, Anni F., Suzanne, Karen, Jason E., Mandy, Raju, Claudine, Shruti, Segun, Katie, and Nandini. You all made the department an enjoyable place to work.

I also want to thank my parental units! P1 and P2, you are the best! Thank you for supporting me emotionally and financially for these past few years. Thank you to my brother, Dr. Casey, for putting it all in perspective. Without you all I would not have succeeded in life.

I also have to thank my neighborhood friends, Skip, Derrick, and Rob for making me laugh and keeping me sane during these final months. Finally, thank you to the recreational facilities at UMd. I swam countless miles in the pool during the last 8 years and countless hours were spent at Cole Field House and the Comcast Center. Thanks to Gary Williams and Juan Dixon as well! Go Terps!!!!

TABLE OF CONENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Amp	ampicillin
bp	base pairs
DIG	digoxigenin
DGI	Disseminated gonococcal infection
ESI	electrospray ionization
Gal	galactose
GalNAc	*N*-Acetylgalactosamine
GC	gas chromatography
Glc	glucose
GlcNAc	*N*-Acetylglucosamine
Hep	heptose
Hex	hexose
Kan	kanamycin
kb	kilobase
Kda	kilodalton
KDO	3-deoxy-D-manno-2-octulosonic acid
LPS	lipopolysaccharide
LOS	lipooligosaccharide
mAb	monoclonal antibody
MALDI-TOF	matrix assisted laser desorption ionization-time of flight
MW	molecular weight
ORF	open reading frame

OS	oligosaccharide
PCR	polymerase chain reaction
PEA	phosphoethanolamine
PID	Pelvic inflammatory disease
Poly-C	poly-cytosine
Poly-G	poly-guanine
PS	polysaccharide
Rham	rhamnose
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
UG	uncomplicated gonorrhea
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

CHAPTER 1. INTRODUCTION

Background and Significicance

a. *Neisseria* and social impact

Every year, 3 billion dollars are spent fighting neisserial diseases (CDC, 2002). The genus colonizes mucosal surfaces including the oropharynx, nasopharynx, and urogenital tracts. *N. gonorrhoeae* and *N. meningitidis* are classified as pathogenic strains of the *Neisseria*. *N. gonorrhoeae* is an obligate human pathogen and the causative agent of gonorrhea. *N. meningitidis* is a common colonist of the nasopharynx. Upon dissemination, meningo-encephalitis can result in 10% mortality (54). Other *Neisseria* are classified as commensal, including *N. sicca*, *N. lactamica*, *N. subflava*, *N. polysaccharea*, *N. cinerea*, *N. mucosa*, *N. flavescens*, and *N. elongata*. The commensals are common colonists of mucosal tracts and generally only cause disease in immuno-compromised individuals (CDC, 2002).

b. Surface structures and generation of neisserial vaccines

Recent studies have shown an increase in the levels of antibiotic resistance in *N. gonorrhoeae* and *N. meningitidis*. For example, gonococcal isolates have been shown to have an increased resistance to penicillin, quinolones, tetracylines, and ciprofloxacin (8). As a result, broader spectrum antibiotics are increasing utilized to fight neisserial infections. These broader drugs are both expensive and have negative side effects on the patient.

Research has turned to the creation of a vaccine to prevent neisserial diseases. Currently, a vaccine is available for *N. meningitidis* strains of serogroups A,C,Y and W135. However, in Europe and America the majority of meningococcal meningitis cases

are caused by strains belonging to serogroup B. No vaccine exists for strains of this serogroup because the capsular polysaccharide is a polymer of α(2-8) linked sialic acid, a structure that is found on human gangliosides. Therefore, a vaccine with these antigens could lead to autoimmunity (37). Additionally, a gonococcal vaccine remains unavailable.

Other strategies in the development of neisserial vaccines focus on locating highly conserved membrane-associated molecules that are normally exposed to the immune system. Lipooligosaccharide (LOS) is one such molecule that is expressed on the neisserial surface and is composed of a lipid portion, the endotoxin, and a carbohydrate component. Vaccines based on LOS focus on the carbohydrate portion, as the endotoxin is toxic. Therefore, significant efforts have been made to identify core regions of LOS that are conserved across the *Neisseria* (132).

c. Structure and synthesis of neisserial LOS

LOS is found on the neisserial surface and is known to play a role in pathogenesis. Most LOS studies have focused on the molecule produced by the pathogenic *Neisseria*, *N. gonorrhoeae* and *N. menigitidis.* As a result, there is a developing picture of the specific LOS structures that these organisms can express and the genes required to produce these molecules. Both the lipid and carbohydrate portions of LOS/LPS have roles in neisserial pathogenesis. Endotoxin activation of TLR 4 receptors on macrophages and endothelial cells stimulates TNF-α and IL1-β, mediators of inflammation (12,29). When overproduced these factors cause septic shock, intravascular coagulation and multiple organ failure. Endotoxin is also responsible for inflammation of the cervix, urethra, and the fallopian tubes in gonococcal infections and

is responsible for the toxicity associated with septicemia in meningococcal infections (5,17). The oligosaccharide (OS) component also plays a part in pathogenesis. First, OS sialyation blocks complement-mediated killing (32,83). Second, specific OS structures such as lacto-N-neotetraose are involved in the attachment to epithelial cells (121). Finally, OS can express structures that mimic antigens found on host cells, protecting the bacteria from the host immune response (84).

The gonococcus and meningococcus have the genetic capacity to synthesize structurally related LOS molecules. Figure 1 displays the composite LOS structure that the pathogenic *Neisseria* are known to produce. LOS is composed of three regions: lipid A, the inner core, and the outer core. The outer core contains three OS chains: α, β, and γ. In order to make LOS, *Neisseria* use genes encoding glycosyl transferases, lipid biosynthesis enzymes, biosynthetic precursors, and machinery to transport molecules to the cell surface. Genes encoding glycosyl transferases are also shown in Figure 1. During assembly, RfaC adds HepI to KDO (124). Next, RfaF adds HepII to HepI (95,109). After these first two steps, α chain synthesis can proceed. LgtF adds glucose to HepI. The rest of the α chain synthesis is directed from a gene cluster, *lgtABCDE* (48). The structure of the α chain is determined by phase variation due to poly-guanine tracts that are contained within *lgtA*, *lgtC*, and *lgtD*. Alterations in the length of the tract can result in the production of a non-functional protein (23,62,140). For example, the premature transcriptional termination of *lgtA* and *lgtC* can truncate the α chain, resulting in a Mab 2-1-L8 reactive molecule. When *lgtD* is in frame, Mab 1B2 reactivity is lost while the binding of Mab 1-1-M occurs. Synthesis of the β chain begins through the addition of Glc to HepII by LgtG (10). *lgtG* contains a poly-cytosine tract that affects the production of

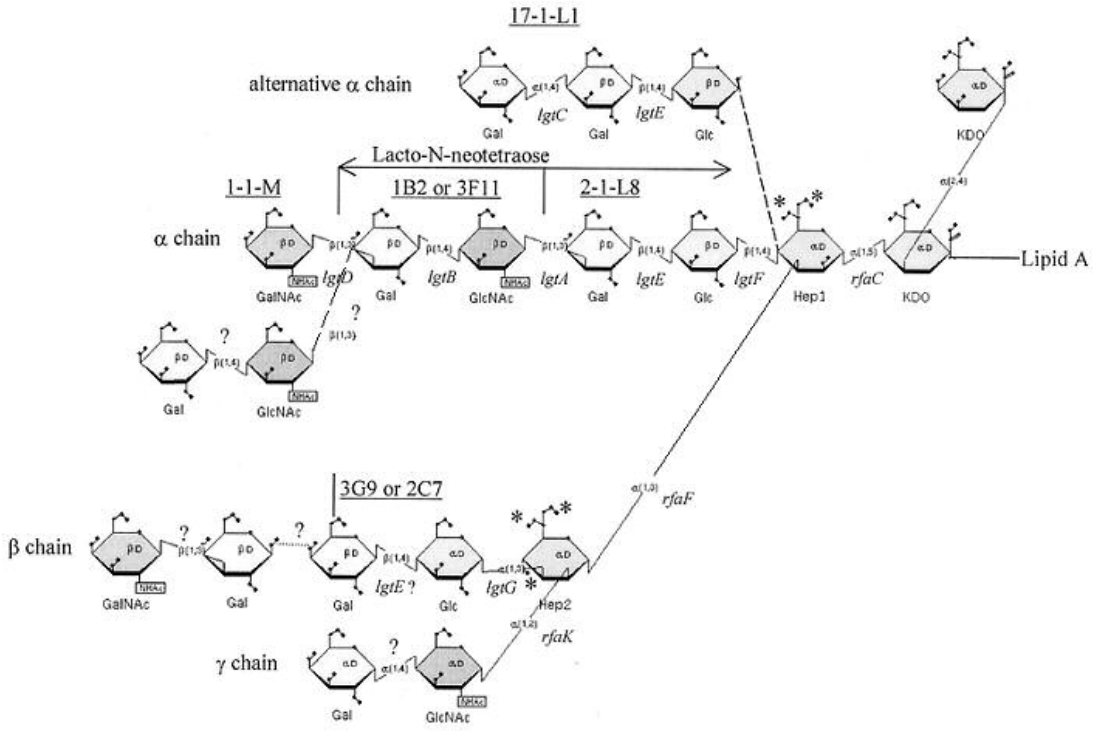


Figure 1. Schematic of the composite neisserial lipooligosaccharide. Genes encoding glycosyl transferases are shown in italics. Unknown enzymes mediating sugar addition to the LOS are represented by question markds. Monoclonal antibodies that bind specific LOS epitopes are underlined. (128).

the protein. When *lgtG* is out of frame, Mab 3G9 and Mab 2C7 reactivity is lost. In this state, PEA may be added to the 3 position of HepII. Finally, RfaK adds GlcNAc to HepII to complete the γ chain (67). In total, 3 loci have been identified that encode glycosyl transferases responsible for LOS synthesis (9,48,68).

d. LOS as a virulence determinant

LOS biosynthesis and phase variation are crucial to neisserial pathogenesis. Phase variation contributes to evasion of the host immune system. For example, the structure recognized by the 1B2 Mab mimics lacto-N-neotetraose of the sphingolipid paragloboside (84,112). The LOS structure recognized by Mab 1-1-M mimics gangliosides, while the alternative α chain mimics globoglycolipids (26). The mimicked structures are normally expressed on host erythrocytes, allowing the bacteria to look like "self" to the host.

Besides molecular mimicry, certain LOS structures are important for host invasion. The lacto-N-neotetraose LOS structure has been the subject of several studies. This molecule interacts with the asialoglycoprotein receptor of primary human urethral epithelial cells (52). This structure also promotes invasion of epithelial monolayers in the absence of Opa proteins (121). Terminal lacto-N-neotetraose producing gonococcal isolates are more often collected from urethritis patients (112). The same structure is also important for serum resistance. Lacto-N-neotetraose can be sialyated using host 5'-cytidinemonophospho-N-acetylneuraminic acid (CMP-NANA) (85). Sialyated strains have a reduced ability to adhere to neutrophils in the absence of complement and decrease oxidative burst in neutrophils (73,104). These structures also result in strains with reduced binding to IgM and IgA (133). Highly pathogenic strains such as *N.*

gonorrhoeae PID2 express terminal lacto-N-neotetraose while commensals generally do not, showing the importance of specific structures for neisserial pathogenicity (6,127).

Individual *Neisseria* are able to simultaneously express different LOS molecules (18). This mechanism allows the Neisseria to survive in different niches within the host. For example, strains of *N. meningitidis* that colonize the nasopharyx and fail to invade have *lgtA* out of frame. However, invasive meningococcal isolates taken from blood have *lgtA* in frame, producing the lacto-N-neotetraose structure (11). LOS phase variation allows for colonizing cells expressing one LOS to alter the LOS to a desired structure that favors survival in different environments.

e. *Neisseria* as pathogens (commensal vs. pathogenic)

Comparatively little is known about the LOS produced by commensal *Neisseria.* For example, *N. sicca* 4320 was responsible for a fatal case of endocarditis in an otherwise healthy adult male. The isolate differed from other *Neisseria* in that it expressed lipopolysaccharide (LPS) (108). The study of the synthesis and structure of this rare isolate is necessary for several reasons. First, the gonococcus and meningococcus may possess genes responsible for the *in-vivo* synthesis of a polysaccharide repeat that increases neisserial virulence. Evidence of this possibility has been shown by the *N. gonorrhoeae* MS11 mkC *in-vivo* addition of polylactosamine to the α chain, shown in Fig.1 (64). *N. sicca* 4320 LPS would serve as a model organism that is able to synthesis the LPS *in-vitro*. Second, *N. sicca* 4320 LPS synthsis genes could be transferred to other *Neisseria*, increasing the virulence of the genus while rendering previously designed neisserial vaccines ineffective.

The second possibility is feasible because there is a well-documented history of genetic transfer between members of the genus. *Neisseria* are naturally competent microbes that preferentially take-up DNA containing the sequence 5'GCCGTCTGAA3' from the environment. The sequence occurs in neisserial chromosomes approximately every 2 kb (28). Since the uptake sequence is located throughout the chromosome, pathogenic and commensal *Neisseria* form a common gene pool where one *Neisseria* can donate DNA to a recipient *Neisseria*. Horizontal genetic exchange is one method by which different *Neisseria* species can acquire new DNA that was originally arranged by vertical gene transfer or obtained by an infrequent non-specific uptake event. As a result, an allele acquired by one *Neisseria* strain has access to the global population of *Neisseria*.

Recent studies show that commensal *Neisseria* serve as a reservoir of genetic diversity and as a result may provide pathogenic strains with the ability to display altered surface antigens. Antigen variability can aide in the evasion of host cell defenses. A striking example comes from a 1980's Meningitis Outbreak in The Gambia, West Africa. 98 serogroup A subgroup IV-I meningococcal isolates were obtained during the outbreak from diseased individuals. Isolates of other subgroups as well as commensal *Neisseria* were collected from healthy persons in the same region. Disease causing isolates contained altered *tbpB* alleles, encoding transferrin binding proteins, many of which were imported from *N. lactamica*. The altered TbpB was selected for within the host (80). Other studies further show horizontal genetic exchange in iron binding genes. Eight meningococcal isolates were found that lacked *hmbR*. These strains contained commensal *Neisseria tbpB*, *lbpB*, and *hpuA* sequences at the *hmbR* locus. This exchange allows for

iron acquisition in pathogenic *Neisseria* by alternative pathways donated by commensal strains (66). In another experiment, sequence analysis was performed to compare *fbp* in commensal and pathogenic *Neisseria*. Nucleotide sequences typical of commensal *fbp* were seen in several strains of pathogenic *Neisseria* (41). Finally, extensive gonococcal (GC) sequence analysis from disease causing isolates demonstrates transfer of *tbpA* alleles within the species. TbpA was altered in loops creating an iron passage pore (21). Horizontal genetic exchange creates diversity to allow *Neisseria* to obtain nutrients to survive in the host environment.

Several studies have examined the other surface structures in the pathogenic *Neisseria* and compared the sequences encoding these proteins to commensal *Neisseria* sequences. The following examples support the notion of frequent horizontal exchange between commensal and pathogenic *Neisseria*, creating a common gene pool. Porins are major components of the outer membrane of the *Neisseria* (137). A 3D model of meningococcal porin was deduced and important regions in its structure were determined. Phylogenetic analysis of these crucial regions revealed interspecies genetic exchange between *N. meningitidis, N. gonorrhoeae, N. lactamica*, and *N. polysaccharea*. An allele of gonococcal, *porB1b* arose by recombination with an allele sharing a common ancestor with *N. lactamica* porin (27). Gonococcal intraspecies transfer of porin genes was studied. 11 *porA* sequences were aligned and synonymous and non-synonymous substitution ratios show evidence of selection for exchanged alleles (120). In another study, *por* sequences of 33 *N. gonorrhoeae* disease-causing isolates were aligned. The sequences encoding the exposed loop structure were transferred between strains. One isolate, MS11 with allele PIB9 actually contained a loop sequence from allele PIB3 (56).

Pilin is another surface structure and is thought to be important for adhesion. PilE has two classes, I and II. Class I are found in the gonococcus and in some meningococcal strains. Class II are found in the meningococcus and commensal strains. The hypervariable disulfide region is antigenically variable and exposed in Class I pilin. Class II pilin lack this region, but contain a 4-10 amino acid insertion that allows diversity of alternate exposed molecules. Recombination can occur between species sharing the same class, causing antigenic variation (2).

The surface molecule Opa is believed to be important for cellular invasion. Commensal *Neisseria* possess genes with significant homology to gonococcal Opa, but those genes are not expressed in commensal strains (137). However, it has recently been shown that commensals express pathogen-like Opa adhesins. These Opas are able to interact with CEACAM1 molecules that were already identified as receptors for pathogenic *Neisseria* (126). Because of the high degree of homology, exchange can occur leading to diversity of Opa expression.

The distribution of *Neisseria* LOS biosynthesis genes have been examined and some are present in some commensal strains (6). The few instances where the genes are present in commensal *Neisseria* demonstrate horizontal gene transfer. Several alleles exist for each *lgt* gene at all 3 loci. One example of *lgt* horizontal gene transfer is at *lgtG*. MC58 contains a 5' *lgtG* sequence that is meningococcal in origin, but the 3' end of the gene is clearly commensal. Also, *N. subflava* 44 has 5' and 3' ends homologous to meningococcal sequences while the internal region is similar to the gonococcus. Other evidence for gene exchange is the similarity of the MC *lgt-1* locus organization with commensal *Neisseria*. For the most part, *lgtC* and *lgtD* are absent from meningococcal

and commensal strains when compared to the gonococcus (141). This evidence demonstrates that interspecies recombination of genes encoding glycosyl transferases occurs. If a new mechanism of LOS/LPS biosynthesis is obtained by a commensal strain, a pathogenic strain has access to that process.

Another advantage that commensal *Neisseria* provide pathogenic *Neisseria* is the ability to become antibiotic resistant. Studies show that commensal strains donate antibiotic resistance genes to the pathogenic species. Meningococcal and gonococcal strains were identified in Spain that were naturally resistant to penicillin (7). Commensal neisserial-derived penicillin binding proteins were responsible for a decreased affinity to penicillin in these strains. Meningococcal strains obtained DNA from *N. polysaccharea* resulting in a mosaic *penA* (92). An analogous case is demonstrated through the investigation of erythromycin resistant gonococci from an outbreak in Uruguay. Commensal *Neisseria* from that region show that *ermF* and *ermB* were acquired in the gonococcus in the early 1990's, although they had been in commensals since 1980 (106). The commensal species are reservoirs of antibiotic resistance genes that are transferred to the pathogenic *Neisseria* by horizontal genetic exchange. All of these examples provide evidence that *Neisseria* are a recombining population with a common gene pool and that commensal neisserial sequences can be acquired by pathogenic *Neisseria*. For this reason, surface structures expressed by the commensal *Neisseria* need to be accounted for during vaccine design.

Specific Aims and Approaches

This study used *N. sicca* 4320 as a model to research the extent to which a non-pathogen differs from a pathogen in order to identify specific traits that increase

neisserial virulence. Strains of *N. sicca* are considered to be commensal, in that they normally colonize mucosal sites and only cause pneumonia and bronchiectasis in immuno-compromised individuals (49). On rare occasions, *N. sicca* colonization has resulted in endocarditis cases in patients that are IV drug users, have heart-valve malformations, or are chronic peritoneal dialysis patients (19,53,91,131). *N. sicca* 4320 was determined to be the causative agent of a fatal endocarditis case in an otherwise healthy adult. This isolate from a normally commensal species represents an increase in virulence and is therefore worthy of study to identify virulence traits that enabled this organism to invade the host and survive in the blood.

Preliminary studies of *N. sicca* 4320 found a surface antigen that was uncharacteristic of the *Neisseria*. Unlike other LOS synthesizing members of the genus, both commensal and pathogenic, 4320 expresses LPS (108). It was suggested that the presence of LPS on the surface of this isolate increased its virulence as another study found that *N. gonorrhoeae* MS11 expressing LOS containing polylactosamine repeats yielded an increased rate of urethritis (64). However, this isolate is unable to synthesize the repeat *in-vitro*, making study of this molecule difficult. Since *N. sicca* 4320 produces LPS *in-vitro* it was thought that study of this molecule might provide insight to the potential ability of pathogenic *Neisseria* to produce polylactosamine repeats *in-vivo*.

The first aim of this study was to identify the polysaccharide structure that is produced by *N. sicca* 4320. Structural analysis was performed through the combination of several methods. MALDI-MS analysis was performed to identify the mass of the OS repeat. GC-MS and lectin binding assays were utilized to determine the sugar composition of the repeat unit. Exoglycosidase digestion and MS^n analysis were

performed to gain information on the connectivity of between the monosaccharides composing the repeat.

Once the structure of *N. sicca* 4320 LPS was known, attention shifted to determining the genetic mechanism responsible for its synthesis. First random sequence analysis of the whole chromosome was done to compare the contents of commensals and pathogenic species. The chromosome of *N. sicca* 4320 was compared to that of *N. menigitidis* MC58 because both normally inhabit the nasopharynx. This analysis was also undertaken to identify potential LPS biosynthetic genes. In order to perform genetic experiments with this organism, a transformation method was developed because 4320 is not naturally competent. Once chemical competent 4320 cells were made, transposon mutagenesis was used to randomly mutagenize the chromosome and other suspected LPS biosynthetic genes. These mutants were screened for 4320 expressing altered LPS through Western blotting with antibody directed against the LPS in hopes of identifying a biosynthetic gene.

Previous studies have found that the presence and specific location of phosphorylation on neisserial LOS plays a role in the level of pathogenicity. For example, a study showed that *N. meningitidis* expressing LOS with phosphoethanolamine (PEA) at the 6 position of HepII were more susceptible to complement mediated killing than strain producing LOS with PEA at the 3 position of HepII (102). Also, it was determined that pathogenic neisseria contain phosphorylcholine on pili whereas commensal species contain the modification on the LOS (115,116). Because phosphorylation of LOS impacts virulence, Southern hybridization experiments were

performed to determine if homologs of genes encoding PEA transferases in pathogenic *Neisseria* are present *N. sicca*.

To further explore the significance of PEA modification of neisserial LOS a second specific aim was developed. This aim was designed to identify and characterize Lpt3, the protein responsible for transferring PEA to the third carbon of HepII (3-HepII) of gonococcal LOS. Previous studies speculated that *lpt3* of the meningococcus was responsible for adding PEA to 3-HepII (81). However, no biochemical evidence was provided to show the transferase function, and it was thought that this gene was not present in *N. gonorrhoeae* FA1090. In this study, a bioinformatic analysis was attempted to show the presence of this gene in the gonococcus. To determine if the protein actually exhibited transferase function at 3-HepII, a gonococcal *lpt3* mutant was constructed. LOS from this strain was isolated and purified and used to react with isolated Lpt3 in a Western blot. Previous studies have shown that the presence of *lpt3* does not necessarily correlate with PEA decoration at 3-HepII of LOS (69,100). Therefore, a phylogenetic analysis of *lpt3* sequences from the pathogenic *Neisseria* was performed to determine if different forms of the enzyme can account for different ratios of PEA decorated LOS. Also, MALDI-MS was used to determine if promiscuity of the PEA transferase enzymes could also contribute to the different PEA ratios.

CHAPTER 2. MATERIALS AND METHODS

Bacterial strains and culture conditions. All strains and plasmids used in this study are described in Table 1. *Neisseria* strains were grown in phosphate-buffered gonococcal medium (Difco) supplemented with 20 mM D-glucose and growth supplements either in broth with the addition of 0.042% $NaHCO_3$ or on agar at 37^{o}C in a CO_2 incubator (136). *E. coli* strains were grown on Luria-Bertani medium (Sambrook, 1989). Kanamycin was used at 50 μg/mL, Ampicillin was used at 60 μg/mL and X-Gal (5-bomo-4-chloro-3-indolyl- -D-galactopyranoside) was used at 35 μg/mL.

Chemicals, reagents and enzymes. Restriction enzymes, glycosidases, and T4 DNA ligase were purchased from New England Biolabs (Beverly, Mass.). All chemicals used for this study were reagent grade or better and were purchased from Sigma Chemical Co. (St. Louis, Mo.) unless otherwise specified in the appropriate materials and methods section. Tris-Tricine gels (16.5%) and running buffer were obtained from Bio-Rad Laboratories (Richmond, Calif.). The Mab 2-1-L8 was graciously provided by Wendell Zollinger (Walter Reed Army Institute of Research, Washington D.C.)

DNA isolation procedures. Chromosomal DNA was isolated using Promega's Wizard^{TM} Genomic DNA Purification Kit. Plasmid DNA was isolated by the Birnboim and Doly alkaline lysis method (14).

Transformation. Competent cells of *E. coli* DH5αMCR and *N. sicca* 4320 were prepared according to Inoue (60). Transformation of *E. coli*

Table 1. Bacterial strains and plasmids used in this study.

Strain or Plasmid	Relevant genotype or phenotype	Source or Reference
Strains		
N. gonorrhoeae F62		P.Frederick Sparling
N. gonorrhoeae F62ΔLgtA	622-bp *BspEI* and *AgeI lgtE* deletion (Mab 2-1-L8 +)	(121)
N. gonorrhoeae F62ΔLgtAlpt3::Tn5	Tn5 insertion at bp 866 of *lpt3* (Mab 2-1-L8 -)	This Study
N. gonorrhoeae F62ΔLgtAΔlpt6	*lpt6* replaced with Spec^r	This Study
N. gonorrhoeae F62ΔLgtA Δlpt6lpt3::Tn5	*lpt6* replaced with Spec^r Tn5 insertion in *lpt3*	This Study
N. gonorrhoeae FA1090	Cervical isolate from DGI	William Shafer
N. gonorrhoeae PID2	Cervical isolate from PID	
N. gonorrhoeae PID305-1	Cervical isolate from PID	(9)
N. gonorrhoeae DGI 5	Blood isolate from DGI	(9)
N. gonorrhoeae PID 032	Urethral isolate from UG	(9)
N. gonorrhoeae DGI Bu	Cervical isolate from DGI	(9)
N. gonorrhoeae PID Ar	Cervical isolate from PID	(9)
N. gonorrhoeae 405 SGI	Urethral isolate from UG	(9)
N. gonorrhoeae 517109	Cervical isolate from UG	(9)
N. gonorrhoeae 659 SGI	Urethral isolate from UG	(9)
N. gonorrhoeae DGI Aj	Synovial fluid isolate from UG	(9)
N. gonorrhoeae 582 SGI	Ureththral isolate from UG	(9)
N. gonorrhoeae 577123	Cervical isolate from UG	(9)
N. gonorrhoeae PID 036	Urethral isolate from UG	(9)

N. gonorrhoeae PID 332	Cervical isolate from PID	(9)
N. gonorrhoeae PID 022	Urethral isolate from UG	(9)
N. gonorrhoeae PID 1	Cervical isolate from PID	(9)
N. gonorrhoeae DGI 11	Urethral isolate from DGI	(9)
N. gonorrhoeae 131 SGI	Urethral isolate from UG	(9)
N. gonorrhoeae 553630 SGI	Cervical isolate from UG	(9)
N. gonorrhoeae PID 22-1	Cervical isolate from PID	(9)
N. gonorrhoeae 553 SGI	Urethral isolate from UG	(9)
N. gonorrhoeae PID 20	Cervical isolate from PID	(9)
N. gonorrhoeae DGI 40	Synovial fluid isolate from DGI	(9)
N. gonorrhoeae 538 SGI	Urethral isolate from UG	(9)
N. gonorrhoeae DGI 8	Blood isolate from DGI	(9)
N. gonorrhoeae DGI 4	Blood isolate from DGI	(9)
N. gonorrhoeae DGI Tc-2	Rectal isolate from DGI	(9)
N. gonorrhoeae DGI 41	Synovial fluid isolate from DGI	(9)
N. gonorrhoeae DGI 48	Cervical isolate from DGI	(9)
N. gonorrhoeae Kb	Urethral isolate from DGI	(9)
N. gonorrhoeae PID 334	Cervical isolate from PID	(9)
N. gonorrhoeae DGI 43	Synovial fluid isolate from DGI	(9)
N. gonorrhoeae PID 335	Cervical isolate from PID	(9)
N. gonorrhoeae PID Au	Cervical isolate from PID	(9)
N. gonorrhoeae PID 305	Cervical isolate from PID	(9)
N. gonorrhoeae PID18	Cervical isolate form PID	(9)

N. gonorrhoeae DGI 13	Blood isolate from DGI	(9)
N. gonorrhoeae 336-1	Cervical isolate from PID	(9)
N. gonorrhoeae 510 SGI	Urethral isolate from UG	(9)
N. gonorrhoeae PID 023	Urethral isolate from UG	(9)
N. gonorrhoeae DGI Tc-1	Synovial fluid isolate from DGI	(9)
N. gonorrhoeae DGI 3	Pharyngeal isolate from DGI	(9)
N. gonorrhoeae 577 SGI	Urethral isolate from SGI	(9)
N. gonorrhoeae PID Br	Cervical isolate from PID	(9)
E. coli DH5αMCR	Cloning host strain F^- mcrA(mrr-hsdRMS-mcrBC)	BRL[a]
M15	Expression strain for vector pQE30	Qiagen
BL21	Expression strain for vector pET15b	Novagen
Plasmids		
pUC19	Cloning vector, replicates in *E. coli* DH5αMCR	BRL[a]
pLPT3	F62 LgtAlpt3 cloned into *NdeI* and *EcoRI* of pUC19	This Study
pET15b	Expression vector, replicates in BL21	Novagen
pET15b::PEA	FA1090 *lpt3* cloned into *NdeI* and *EcoRI* of pET15b	This Study
pQUE30	Expression vector, replicates in M15	Qiagen
pQE30::PEA	FA1090 *lpt3* cloned into *BamHI* and *PstI* of pQUE30	This Study

[a]Bethesda Research Laboratories is now part of Invitrogen

and *N. sicca* 4320 was done according to the heat-shock protocol (Sambrook, 1989). DNA transformation into *N. gonorrhoeae* was done by resuspending T1 cells in GCP broth containing 1x Kellogg's solution, 0.042% $NaHCO_3$, 10 mM $MgCl_2$, and 1-2 μg of the DNA of interest (122). Cells were incubated for about 3 h with shaking at 37 ^oC. Cells were plated onto GCK plates containing the selective antibiotic. The spot transformation consisted of suspending 10 isolated T1 colonies in 500 μl GCP broth containing 1% 1M $MgCl_2$. Serial dilutions of cells were spotted onto GCK plates and 2 μg of DNA was mixed into each spot. Plates were allowed to dry prior to overnight incubation. Colonies were patch plated onto GCK and GCK plates containing kanamycin (50).

LOS/LPS extraction. Quick preparations of gonococcal LOS were prepared from plated cultures as described by Hitchcock and Brown (55). Cells were swabbed from GCK agar plates and resuspended in 3 mL of 100mM PBS (pH 7.5) to a Klett of 100. An aliquot was centrifuged at 10,000 rpm and the pellet re-suspended in 50 μl of lysing buffer. Proteinase K (10 μl of 1 mg/mL stock) was added and the sample was heated for 1 hour at 60^{o}C. While still hot, samples were diluted (1:25 in lysing buffer to observe LOS and 1:5 to view the O-Repeat portion of LPS). For SDS-PAGE analysis, samples were boiled for 10 min and 5 μl was loaded onto the gel. Purified LOS was obtained using the Westphal and Jann hot-phenol method with the following modifications (135). To collect approximately 50 mg of LOS/LPS, 16 L of *Neisseria* were grown in broth overnight. Cells were pelleted at 5,000 rpm and re-suspended in 68^{o}C HPLC water (25 mL for each 1 L of overnight culture). An equal volume of 68^{o}C phenol was added and the solution

was periodically vortexed for 30 min while remaining hot. The mixture was centrifuged at 5,000 rpm and the aqueous layer was collected. The remaining phenol layer was extracted with HPLC water, and the aqueous layers were combined. Crude LOS/LPS was precipitated by adding two volumes of cold acetone, sodium acetate (200 mg / 100 mL), and storing at –80^{o}C overnight. After thawing on ice, crude LOS/LPS was centrifuged and the pellet was re-suspended in HPLC water. DNAse and RNAse (1000 units each) were added and after a 1 hr bench top incubation, hot phenol-water extractions were performed again. After acetone precipitation, LOS/LPS was dialyzed against several changes of HPLC water and analyzed for purity using the Wavescan program on the spectrometer. LOS/LPS underwent ultra-centrifugation at 30,000 rpm for 3 hrs and the jelly like pellet was re-suspended in a minimal amount of HPLC water, frozen and lyophilized to a powder.

SDS-PAGE analysis. LOS/LPS suspension was boiled for 10 minutes immediately before loading. Approximately 0.1 μg of LOS or 0.5 μg LPS was subjected to SDS-PAGE on a 16.5% Tris-Tricine gel (from Bio-Rad) in Tris-Tricine running buffer at a constant current of 30 mA for 3 hrs. The gel was fixed overnight in 40% ethanol-5% acetic acid. The next day, the LOS/LPS was visualized by silver staining. The gel was oxidized in 0.833% periodic acid for 5 min. The gel was washed with HPLC water every 30 min for two hrs. Silver stain (0.0225 M NaOH, 0.42% NH_4OH, 0.047 M $AgNO_3$) was added for 15 min. The gel was rinsed again with HPLC water for 45 minutes, changing the water every 15 min. LOS/LPS bands were developed by incubation in developing solution (0.005% citric acid, 0.007% formaldehyde).

Western Blot analysis. After SDS-PAGE, LOS/LPS molecules were electrotransferred onto Immobilon-P membrane (Millipore Corp.) in a Tris-Tricine-methanol buffer (10 mM Tris [pH 8.3], 10 mM Tricine, 0.01% SDS, 20% methanol) at a constant voltage of 100 V for 1 h following the protocol provided by Bio-Rad Corp. After air drying for 1 h, the membrane was processed in buffer (100 mM Na_2HPO4 (pH 7.5), 2% casein, 0.2% NaN_3, 0.002% phenol red) to block all nonspecific binding sites. The membrane was washed three times with TBS (50 mM Tris, 150 mM NaCl, pH 8.0) and incubated overnight with primary Ab. After washing away the unbound Ab with TBS, the remaining Ab was detected by reacting the membrane with horseradish peroxidase-labeled secondary Ab for two hours, washing with TBS, and testing for the presence of bound horseradish peroxidase using the following development conditions (50 mM Tris [pH 8.0], 0.006% H_2O_2, 0.08% 4-chloro-1-napthol). When colony blot analysis was performed, overnight colonies were transferred to a nitrocellulose membrane (Schleicher and Schuell, Keene, N.H.) and processed as described above.

PCR. Primers used in this study are listed in Table 2. The polymerase chain reaction was performed using the Expand Long Template PCR kit (Roche) according to the manufacturer's specifications. Primers were purchased from Integrated DNA technologies (Coralville, Iowa). PCR reactions were resolved on 1% agarose gels containing 500 μg/ml of ethidium bromide in Tris-Borate EDTA running buffer. Sequencing of PCR products was performed by Macrogen Inc. (Seoul, South Korea).

Table 2. Primers used in this study.

Primer	Sequence (5'- 3')	Description
PTE-6	ATTCCAAGGATCCAAAAAATCCCTTTT CGTTCTCTTTCTGTATTC	Amplifies the 5' end of *lpt3* with *BamHI* site
PTE-7	CGTACTGCAGATCTAAATTTGTCTTTA TTGGTTTTCACGCTGCTTATT	Amplifies the 3' end of *lpt3* with *PstI* site
PTE-4	ATTCCAACATATGAAAAAATCCCTTTT CGTTCTCTTTCTGTATTC	Amplifies the 5' end of *lpt3* with *NdeI* site
PTE-5	CGTAGAATTCATCTAAATTTGTCTTTA TTGGTTTTCACGCTGCTTATT	Amplifies the 3' end of *lpt3* with *EcoRI* site
PEA-R	CCGGAATTCCGGAATCACCAATATTT CCGCCATGCAGTCGTG	Amplifies a sequence downstream of *lpt3* with *EcoRI* site
PEA 1	TAGCCGTCTTTTTACGCAGAA	Amplifies 5' end of *lpt3*
PEA 2	CGGCAACCGCACGCCATAGCG	Amplifies 3' end of *lpt3*
PEA 3	ACGGACGCGAAACTTCGCCGT	Amplifies the middle of *lpt3*
PEA 4	GCGCATTGTTGCAGCCTCAAG	Amplifies the middle of *lpt3*
EZTN-F	TCTTGCCATCCTATGGAACTGCCTCGGTGA	Amplifies the sequence downstream of TN5
Kan-R	GGGCAATTGCTGAAGCTTGCATGCCTGCAGG	Amplifies 3' end of Kan^r with *MfeI* site
Kan-F	GGGGAATTCAACCATCATCGATGAATTGT	Amplifies 5' end of Kan^r with *EcoRI* site
Lpt-6-F	ATGGTTGCCTATGCTTTCCTATTTTTGTT TGTAACGGCGG	Amplifies 5' end of *lpt6*
Lpt-6-R	AACGGGCAATTTTCAAAACGTCGTACA GATAGTTGCGGTA	Amplifies 3' end of *lpt6*
O6-PEAR	ATTTGAATTCGGCATTCCCGAATACG	Amplifies 3' end of *lpt6* region with *EcoRI* site
O6-PEAF	CGAAGCTTTCCCAGCCGCTTTGG	Amplifies 5' end of *lpt6* region with *HindIII* site
Lpt6 MutF	CCCCTGCAGTAGCTTTTTGCCGCACGGC	Mutation primer to replace *lpt6* with Spec^r cassette with *PstI* site

Lpt6 MutR	CCCCTGCAGCCCGCCCCACTCCTCAAA	Mutation primer to replace *lpt6* with Spec^r cassette with *PstI* site
rfbB-F	CCGGAATTCATGCAAACCGAAGGCAA AAAAAACAT	Amplifies 5' end of *rfbB* with *EcoRI* site
rfbB-R	CCGGAATTCTATGCGTTTTGCCGCCGG GTTTTGTT	Amplifies 3' end of *rfbB* with *EcoRI* site
rfbA-F	CCGGAATTCGAAAAGATGAAAGGCAT CATACTGGC	Amplifies 5' end of *rfbA* with *EcoRI* site
rfbA-R	CCGGAATTCATTTGCCGATTAGGTGCA GCAGGTAT	Amplifies 3' end of *rfbA* with *EcoRI* site
rfbC-F	CCGGAATTCATGGACATCATCGACAC CGCCCTCCC	Amplifies 5' end of *rfbA* with *EcoRI* site
rfbC-R	CCGGAATTCTCATCAGGCGGTAAACT TGCCCTGCC	Amplifies 3' end of *rfbA* with *EcoRI* site
rfbB-F-2-up	CCGGAATTCGGCCGTCTGAAATGCAA ACCGAAGGCAAAAAAAAACAT	Amplifies 3' end of *rfbA* with *EcoRI* site and Neisserial Uptake Sequence

The underlined sequences represent the location of the restriction enzyme site

Isolation and Purification of Lpt3. Recombinant DNAs were transformed into strain M15 (vector pQE30) or BL21 (vector pET15b). Inoculated cultures were grown at 37^{o}C until the OD 600 reached a value between 0.4 and 0.6. IPTG inducer was added and cultures were incubated at 37^{o}C for 4.5 hrs. The cultures were sonicated and proteins were purified on Ni-agarose columns according to the manufacturer's specifications (Qiagen, Valencia, CA) except that after loading the materials, the column was washed with sonication buffer (10 volumes of column), followed by washing buffer for 16 hrs. Proteins were eluted from the column with increasing concentrations of imidazol (2 X 1 mL 5 mM; 2 X 1 mL 10 mM; 2 X 1 mL 25 mM). All buffers contained 1 mM PMSF.

Enzymatic Reaction. Reactions contained 3μl of 10 X concentrated buffer (500 mM Tris-HCl, pH 8.8; 50 mM β-mercaptoethanol; 1 % Trition X-100), 3 μl of $MgCl_2$ (100mM), 2 μl of 1-palmytiol-2-oleyol-phosphatidylethanolamine (Fc 1 mM), and either 3 μl of LOS (1 mg/ml) or cell suspension to a final volume of 16 μl. Prior to initiating the reaction, phosphatidylethanolamine was dissolved a 1:4 mixture of methanol and chloroform and stored at –80^{o}C. To begin the reaction, the phosphatidylethanolamine solution was added to an eppendorf tube and dried in a CO_2 incubator. The other components were added into the reaction tube and mixed until the PEA donor was suspended. Lastly, 1-2 μg of enzyme was added and the reaction was carried out for 90 min at 30^{o}C.

MALDI-TOF MS Analysis. Purified LOS/LPS was O-deacylated prior to MALDI-MS (64). Anhydrous hydrazine (20 μl) was added to 0.5 mg of LOS and kept at 37^{o}C with

periodic vortexing for 20 min. O-deacylated LOS was precipitated with –20°C acetone and centrifuged at 12,000 X g for 20 min. The pellet was washed with cold acetone and re-suspended in H_2O at 2 μg/μl. Samples were desalted with cation exchange beads (Dowex 50X) and combined with 100 mM 2, 5-dihydrobenzoic acid concentrated in MeOH. Negative ion MALDI-MS was performed in linear mode with delayed extraction on a Voyager Elite TOF instrument equipped with a 337 nm nitrogen laser (PerSeptive Biosystems, Framingham, MA). Analyses were performed with a 150-ns time delay and a grid voltage of 92-94% of full acceleration voltage, 20 kv, and external calibration.

ESI-MS Analysis. 500 μg of each sample of purified LOS/LPS were combined with 1 μg of β-cyclodextrin standard. Acetic acid in water (200 μl of 1%) was added to the samples and these contents were heated at 100°C for 1 hr to remove the glycan from the lipid. After centrifugation, the liquid layers were saved and dried by vacuum centrifugation. Samples were derivatized with a pyrazole reducing-end protecting tag by adding 20 μl of anhydrous hydrazine to the dried sample. After vacuum centrifugation, 50 μl of 10% 2,4-pentanedione in water was added and the cyclic pyrazole protecting group was formed during vacuum centrifugation. Samples were methylated, dried, and reconstituted in MeOH (89).

MS^n Analysis. Samples (500 μg) of purified LOS/LPS were combined with 1 μg of β-cyclodextrin standard. Acetic acid in water (200 μl of 1%) was added to the samples and these contents were heated at 100°C for 1 hr to remove the glycan from the lipid. After centrifugation, the liquid layers were saved and dried by vacuum centrifugation. Samples

were reconstituted in 100 µl MeOH and analyzed by NSI-MS/MS^n Finnigan LTQ Classic with Nanospray Ionization Source with a flow rate of 0.5µl/min and 1.3kV spray voltage.

Glycosidase Digestion of O-Repeat. 600 ng of purified *N. sicca* 4320 LOS and O-Repeat was digested overnight with β-N-Acetyl-hexosaminidase at 37^{o}C. Dilutions of this reaction along with identical amounts of undigested *N. sicca* 4320 LOS and O-Repeat were subjected to SDS-PAGE and silver staining.

GC-MS of O-Repeat. HCl (1N) in anhydrous MeOH was added to dried samples of *N. sicca* 4320 LOS and O-Repeat, L-Rhamnose, L-Fucose, and D-GlcNAc. These samples were flushed with N_2 prior to incubation at 80^{o}C for 16-24 hrs. After evaporating to dryness under N_2 at 35-40^{o}C, MeOH was added to eliminate HCl. Then MeOH (200 µl), pyridine (20 µl), and acetic anhydride (20 µl) were added and tubes were vortexed then held at room temperature for 20 min. Samples were evaporated to dryness under N_2 at 35-40^{o}C before adding toluene and acetic acid in excess acetic anhydride. To the methylated glycosides Tri-Sil (200 µl) was added and the tubes were flushed with N_2 and placed at 80^{o}C for 20-30 min. After rapidly cooling tubes to room temperature tubes were again evaporated to dryness under N_2 at 35-40^{o}C. The remaining white residue was washed with two 100 µl washes of n-hexane. The combined washes were used for GC-MS analysis.

Southern Hybridization. Analysis was performed following a combination a protocols from Sambrook and Roche Molecular Biochemicals. Chromosomal DNA was digested

overnight and separated on a 1% TBE agarose gel by electrophoresis. DNA was transferred by capillary action to a pre-wetted positively charged nylon membrane (Roche, Indianapolis IN) using alkaline transfer buffer (0.4M NaOH, 1M NaCl). After overnight transfer, the membrane was removed, air dried, and UV-crosslinked for 5 minutes. The membrane was neutralized while shaking in buffer (0.5M Tris-HCl pH 7.2, 1M NaCl) for 15 min. The membrane was pre-hybridized for 1 hr at 60^{o}C in 20 mL solution (5X SSC, 2% dry milk blocking, 0.1% N laurylsarcosine, 0.02% SDS) for every 100 cm^{2} of membrane. After 1 hr, denatured probe (100 ng of random primed DIG labeled PCR product) was added and hybridization proceeded overnight at 60^{o}C. After hybridization 4 stringency washes were performed. Wash 1 (1X SSC, 0.1% SDS) was done twice at room temperature for 5 min and Wash 2 (0.1X SSC, 0.1% SDS) was performed twice at 60^{o}C for 15 min. The membrane was equilibrated in Buffer A (100 mM Tris-HCl, 150 mM NaCl, pH 7.5) for 1 min. The membrane was placed in blocking solution consisting of 1 g of dry milk in 100 mL of Maleic Acid Buffer (0.1M Maleic Acid, 0.15M NaCl, pH 7.5) for 1 hr. Anti-DIG alkaline phosphatase Ab was added to the blocking solution at a 1 : 20,000 dilution for 30 minutes. Washing was performed in buffer (0.1M Maleic Acid, 0.15M NaCl, 0.3% Tween 20, pH 7.5) twice for 15 min each. Detection buffer (0.1M Tris-HCl, 0.1M NaCl, pH 9.5) was added to the membrane for 5 min. After removal of the detection buffer the membrane was placed in a clear plastic bag and the substrate CSPD was placed on the membrane for 5 min. CSPD was removed and the membrane was incubated in the bag at 37^{o}C for 15 min. The membrane was exposed to X-ray film for different lengths of time.

Transposition Reaction. Transposon mutagenesis was performed on both plasmid and chromosomal DNA using Epicentre's EZ::TN^{TM}<Kan-2>Insertion Kit (Madison, Wisconsin) (47,105). 0.2 μg of DNA was reacted with an equal molar amount of transposon in the presence of transposase for two hours at 37^{o}C. For neisserial transformations, the entire reaction was used. For *E. coli* transformations, 2 μl of the reaction was used.

Table 3. Media compositions/liter H_2O

LB broth (pH 7.0)	
LB broth, miller (Difco)	40 g
LB agar	
LB agar, miller (Difco)	40 g
GCP broth (pH 7.0)	
Protease peptone #3	15 g
Soluble starch (Difco)	1 g
KH_2PO_4	1 g
K_2HPO_4	4 g
NaCl	5 g
GCK agar	
Gonococcal base medium (Difco)	36 g
Bacto-agar (Difco)	5 g
Kellogg's supplement[a]	10 mL
Kellogg's supplement[b]	
Glucose[c]	400 g
Glutamine	5 g
Ferric nitrate	0.5 g
Thiamine pyrophosphate	20 mg

[a] Added after media is autoclaved and cooled prior to pouring
[b] Filter sterilized
[c] Need to be added slowly during stirring

Table 4. Solution compositions

Soultions	Composition
<u>Buffers</u>	
TBE buffer	89 mM Tris-HCl (pH 8.0), 89 mM boric acid, 2 mM EDTA
TAE buffer	40 mM tris-acetate (pH 8.0), 2 mM EDTA
GTE buffer	25 mM Tris-HCl (pH 8.0), 50 mM glucose, 10 mM EDTA
SSC buffer	0.15 M NaCl, 15 mM Na_3Citrate
<u>Southern blot solutions</u>	
Transfer buffer	1.5 M NaCl, 0.5 M NaOH
Neutralization buffer	0.5 M Tris-HCl pH 7.2, 1 M NaCl
Hybridization buffer	5 X SSC, 1% dry milk blocking reagent, 0.02% SDS, 0.1% *N*-lauroylsarcosine
High stringency wash	2 X SSC, 0.1% SDS
Low stringency wash	0.5 X SSC, 0.1% SDS
Maleic acid buffer	0.1M maleic acid, 0.15M NaCl; adjusted to pH 7.5
Blocking stock solution	10% (w/v) dry milk in maleic acid buffer
Blocking buffer	Dilute the blocking stock 1:10 in maleic acid buffer
Washing buffer	Maleic acid buffer plus 0.3% (v/V) Tween20
Detection buffer	0.1M Tris-HCl, 0.1 M NaCl, pH 9.5
<u>LOS analysis solutions</u>	
Filler solution	100 mM Na_2HPO4 (pH 7.5), 2% casein, 0.2% NaN_3, 0.002% phenol red
TBS	50 mM Tris, 150 mM NaCl, pH 8.0
Lysing solution	1M Tris-HCl (pH 6.8), 2% SDS, 4% β-mercaptoethanol, 10% glycerol, bromophenol blue to saturation
Fixing solution	5% acetic acid, 40% ethanol
Oxidizing solution	0.83% periodic acid
Silver stain	0.0225 M NaOH, 0.42% NH_4OH, 0.047 M $AgNO_3$
Developing solution	0.005% citric acid, 0.0007 % formaldehyde
Western transfer solution	10 mM Tris [pH 8.3], 10 mM Tricine, 0.01% SDS, 20% methanol

CHAPTER 3. BIOINFORMATIC SEQUENCE ANALYSIS OF THE *N. SICCA* 4320 CHROMOSOME

INTRODUCTION

The *Neisseria* genus is composed of genetically related species that cause different disease outcomes in the human host (58). For example, colonization with *N. gonorrhoeae* results in patients that experience different degrees of gonococcal infection; disseminated gonococcal infection (DGI), pelvic inflammatory disease (PID) or uncomplicated gonorrhea (UG). Colonization of the nasopharynx with *N. meningitidis* is common in the population. Dissemination into the bloodstream can occur, allowing the meningococcus to access and cross the blood-brain barrier to cause meningitidis. Commensal *Neisseria* are also typically found in the nasopharynx. However, these species do not typically cause disease.

With the availability of complete genome sequences of *N. meningitidis* MC58 and Z2491 and *N. gonorrhoeae* FA1090, several studies have examined the chromosomes from these strains to identify similarities and differences. One study compared the gonococcal and meningococcal chromosomes by subtractive hybridization to identify meningococcal sequences that might encode proteins that are necessary to cause bacteremia and invasion of the meninges. One island was identified, encoding a putative siderophore receptor and a disulfide oxidoreductase, and shown to be important for the meningococcus to achieve high levels of bacteremia (74). Another study used DNA array technology to identify conserved sequences from the pathogenic *Neisseria* strains that are not present in the commensal *N. lactamica*. Genes specific to *N. meningtidis* encoded capsule proteins, Frp toxin, superoxide dismutase, hemagglutinin, and a hemolysin secretion system. Pathogen specific genes encoded pilin adhesion proteins,

immunoglobin protease, hemoglobin receptor, outer membrane peptidase, siderophore receptor, porin, opacity proteins, and hemolysin. These genes are not located on pathogenicity islands, but rather scattered throughout the chromosomes (94). Whole genome comparisons have also been performed on a large collection of meningococcal strains. Chromosomes of meningococcal isolates with pathogenic potential were found to contain a prophage that excises from the chromosome and is secreted by the meningococcus. Because the phage is not found in phylogenetic groups that typically do not invade, it is hypothesized that the secreted elements play a major role in the host invasion (13). These studies show that whole genome analysis can be used to identify neisserial virulence determinants and that that these factors can either be at various chromosomal loci or located within defined regions.

In this study, a bioinformatic approach was used to analyze the *N. sicca* 4320 chromosome for several reasons. First, a comparison of pathogenic *Neisseria* and non-pathogenic *Neisseria* chromosomal organizations has yet to be reported. Study of the organizational differences amongst the neisserial classes can explain why some members of the genus cause disease while others maintain commensal relationships with the host. Second, *N. sicca* 4320 produces LPS which is thought contribute to its pathogencity. Other gram-negative bacteria such as *E. coli*, *Shigella*, and *Salmonella* contain genes that encode LPS biosynthesis proteins that are located on a genetic island (71,134). Although the order of genes varies, all of the islands contain genes for the synthesis of nucleotide sugar precursors, the transfer of sugars to build the O-unit, and the machinery necessary to carry out assembly and processing steps. This *rfb* locus has been shown to be mobile in that it is associated with insertion elements and a low G+C content compared to flanking

housekeeping genes (103,123,134). Therefore, it is possible that *N. sicca* 4320 LPS might be encoded by a homologous island and 4320 chromosomal analysis could allow for the identification of *rfb* homologs.

RESULTS

Generation of *N. sicca* 4320 chromosomal EcoRI and AccI fragments.

Bioinformatic analysis of *N. sicca* 4320 was performed by obtaining random chromosomal sequences for comparison against the *N. meningitidis* MC58 chromosome. Chromosomal DNA from *N. sicca* 4320 was isolated in accordance with Promega's Wizard™ Genomic DNA Purification Kit. Chromosomal DNA was digested with EcoRI yielding chromosomal fragments ranging from 600 bp to greater than 9 kb upon agarose gel electrophoresis. Collectively, the chromosomal fragments were ligated with the pUC19 cloning vector at the EcoRI site. To collect the shotgun clones, the ligation reaction was transformed into *E. coli* DH5αMCR cells. The transformation reaction was applied to LB plates containing ampicillin (60 μg/mL) and X-gal (35 μg/mL) in order to select for transformed cells and screen for clones containing the random chromosomal insertions.

White *E. coli* transformants were collected from the transformation plates because they likely contained the 4320 random clones. Plasmids were prepared from individual white colonies to verify that the clones were present and contained random chromosomal insertions. As shown in Fig. 2A, digestion of the random clones with EcoRI showed that 4320 chromosomal fragments of varying sizes were ligated into pUC19. Upon digestion with EcoRI, clones containing inserts yield two different size products, a 2.7 kb band and a varying band that is the actual insert. Therefore, this technique was useful to obtain a random sampling of the chromosome. An analogous process was performed with the enzyme AccI, that also yielded random insertions into pUC19, as depicted in Fig. 2B.

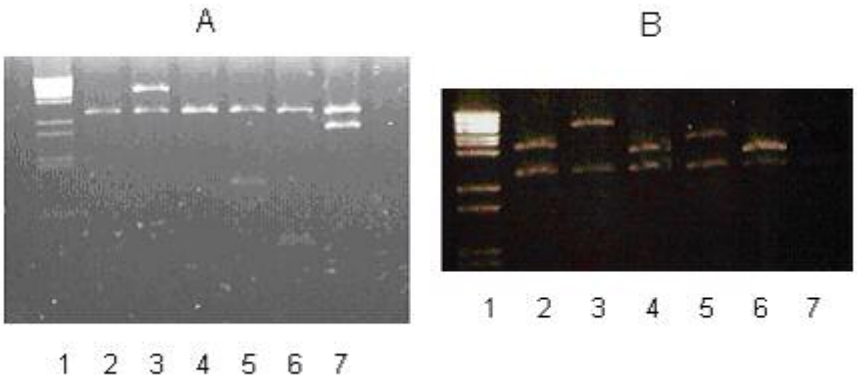


Figure 2. Digestions of random *N. sicca* 4320 shotgun clones. Figure 2A shows EcoRI digestions of random *N. sicca* 4320 shotgun clones. Lane 1 shows the λ BstEII standard. Lane 2 contains pUC19 digested with EcoRI. Lanes 3 through 7 show EcoRI digested plasmids that were extracted from individual white ampicillin resistant *E. coli* transformants. Figure 2B shows AccI digestions of random *N. sicca* 4320 shotgun clones. Lane 1 shows the λ BstEII standard. Lane 7 contains pUC19 digested with AccI. Lanes 2 through 6 show AccI digested plasmids that were extracted from individual white ampicillin resistant *E. coli* transformants.

Sequence Homologies of *N. sicca* 4320 EcoRI and AccI Chromosomal Fragments.

Plasmid DNA was invidually prepared from random isolates and the nucleotide sequence of each insert was determined by Macrogen (Seoul, Korea). Sequencing primer M13-F was selected because it hybridizes to a region upstream of the pUC19 multiple cloning site that is present in each of the constructed clones regardless if the clone was manipulated at the EcoRI or AccI site. This primer allowed for the sequencing of one end of each of the inserts. The typical sequence length obtained for each clone was approximately 850 bp. These sequences of 188 clones were used as query sequences for both the BlastN and tBLASTx databases so that the likely protein encoded by the sequence could be determined as well as the extent to which the nucleotide sequences encoding the homologous proteins are identical. The results of the BLAST searches of sequences from the EcoRI and AccI clones are respectively shown in Tables 5 and 6.

When compared to the *N. menigitidis* MC58 chromosome, the average nucleotide identity of the *N. sicca* 4320 fragments is greater than 80%, indicating the close relatedness between the two organisms. The 4320 fragments showed homology to nucleotides encoding several classes of proteins in the meningococcus. For example, sequences that encode DNA synthesis proteins were identified, specifically those composing the chi, gamma, and tau polymerase subunits. Sequences that encode tRNA for protein synthesis were identified such as, aspartyl-tRNA synthetase and lysyl-tRNA synthetase. Sequences were found that encode proteins involved in cellular respiration such as, acetyl-CoA carboxylase, glucose-6-phosphate isomerase, fructose-bisphosphate aldolase, cytochrome c-type biogenesis protein, and ATP synthase F_0 C subunit.

Sequences were also identified that are predicted to encode proteins necessary for cell division as well as nutrient transport.

Besides the identification of DNA sequences that encode proteins required for cellular metabolism, *N. sicca* 4320 contains genetic material to synthesize membrane proteins that are typically associated with neisserial pathogenesis. For example, pathogenic *Neisseria* contain proteins to acquire and utilize iron (43). 4320 contains the nucleotides sequences that are predicted to encode ferredoxin, an iron regulated outer membrane protein, a hemagglutin/hemolysin related protein, and a hemolysin activation protein. Pilin has been shown to play a role in the attachment of *Neisseria* to host tissues (90). The 4320 sequence analysis shows that DNA contains information to encode components of the pilus apparatus, PilQ and PilM as well as a pilin glycosylation protein. In addition, 4320 contains DNA that may encode LOS synthesis proteins. For example, acyl-(acyl-carrier-protein)—UDP-N-acetylglucosamine O-acyltransferase (LpxA) and UDP-3-O-3-hydroxymyristoyl N-acetylglucosamine deacetylase are likely involved in the synthesis of the lipid A component of LOS. The 4320 chromosome contains sequences with 56% and 59% nucleotide identity to MC58 glycosyl transferase genes *lgtB* and *lgtG*. The presence of these predicted LOS biosynthesis genes is consistent with the fact that commensal *Neisseria* produce LOS, yet many of these transferase genes remain undefined because of low or no homology to the known pathogenic *Neisseria* LOS synthesis genes (6).

In addition to sequences that showed homology to regions of the MC58 chromosome, 4320 fragments were generated that failed to show homology to any genes found in *N. meningitidis*. Studying this group of sequences is important to gain insight on

the differences between a commensal and a pathogenic chromosome. Also, because 4320 LPS synthesis is unique to this neisserial strain, this category of sequences could provide clues to the biosynthetic mechanism of this molecule. Only five of the approximately 180 analyzed clones contained inserts that showed no homology to MC58 chromosomal DNA. 4320 fragments in this sequence category were homologous to nucleotide sequences that encode the following, a glycine rich structural cell wall protein of *Arabidopsis*, an endo-1,4-beta-glucan branching enzyme of *Xylella fastidiosa*, a type II restriction enzyme of *Haemophilus parainfluenzae*, a pyruvate formate lyase activating enzyme of *Yersinia pestis* KIM and a 4-oxalocrotonate decarboxylase of *Acinetobacter*. The sequence fragment that showed homology to the glucan branching enzyme is significant because this enzyme is involved in polysaccharide biosynthesis. Therefore, this region needs to be considered as a candidate for encoding proteins necessary for LPS biosynthesis. The presence of a restriction enzyme that is similar to that encoded by *Haemophilus* is significant because the *Neisseria* and *Haemophilus* both inhabit the same niches on the human host, allowing for exchange of genetic material. Also the fragments reveal that *N. sicca* 4320 may differ in metabolic activity compared to *N. meningitidis* MC58, as indicated by the potential presence of pyruvate formate lyase and 4-oxalocrotonate.

Table 5. Homology of *N. sicca* 4320 to *N. meningitidis* MC58 (EcoRI)

Clone	Homology To Strain	Protein Function	%-N	MC58 Pos.
116	*Neisseria meningitidis* MC58	hypothetical protein & translation elongation factor EF-P	89%	951197-952074
125	*Neisseria meningitidis* MC58	phosphoenolpyruvate synthase	90%	649925-650757
126	*Neisseria meningitidis* MC58	malate oxidoreductase (NAD)	89%	699651-700484
127	*Neisseria meningitidis* MC58	gamma-glutamyltranspeptidase	86%	2081758-2081967
129	*Neisseria meningitidis* MC58	adenylosuccinate lyase	86%	291247-291868
137	*Neisseria meningitidis* MC58	tetrapac protein & putative colicin V production protein	80%	718294-718858
138	*Neisseria meningitidis* MC58	conserved hypothetical protein	81%	1500488-1501061
144	*Neisseria meningitidis* MC58	conserved hypothetical protein	84%	1500511-1501073
145	*Neisseria meningitidis* MC58	probable nucleotide-binding protein	84%	1324229-1324919
149	*Neisseria meningitidis* MC58	elongation factor P (EF-P)	89%	951514-952077
152	*Neisseria meningitidis* MC58	putative nitrogen regulation protein (NtrY)	79%	122422-123288
157	*Neisseria meningitidis* MC58	putative cytochrome c-type biogenesis protein	80%	1893352-1893872
161	*Neisseria meningitidis* MC58	DNA polymerase III, alpha subunit	87%	1921364-1921860
188	*Neisseria meningitidis* MC58	valyl-tRNA synthetase	92%	170097-170884
208	*Neisseria meningitidis* MC58	dihydroxy-acid dehydratase	90%	1156170-1156925
193	*Neisseria meningitidis* MC58	imidazoleglycerol-phosphate dehydratase	96%	1135026-1135476
402	*Neisseria meningitidis* MC58	UDP-3-O-3-hydroxymyristoyl N-acetylglucosamine deacetylase	97%	15445-15705
405	*Neisseria meningitidis* MC58	adenylosuccinate synthetase	84%	841119-841485
406	*Yersinia pestis KIM*	pyruvate formate lyase activating enzyme 1		
407	*Neisseria meningitidis* MC58	ferredoxin, 2Fe-2S type (fdx-1)	60%	1144525-1144687
416	*Neisseria meningitidis* MC58	probable aminomethyltransferase (Glycine cleavage system T protein)	94%	602502-602974
418	*Neisseria meningitidis* MC58	translation elongation factor EF-G	92%	148867-149035
420	*Neisseria meningitidis* MC58	acyl-(acyl-carrier-protein)--UDP-N-acetylglucosamine O-acyltransferase (LpxA)	91%	175430-175906
425	*Neisseria meningitidis* MC58	Phosphoribosylformylglycinamidine synthase (FGAM synthase)	95%	2105024-2105505
431	*Neisseria meningitidis* MC58	adenylosuccinate lyase	93%	291791-291989
438	*Neisseria meningitidis* MC58	nitrate/nitrite sensor protein NarX	83%	1255173-1255728
439	*Neisseria meningitidis* MC58	conserved hypothetical protein	92%	328807-328983
443	*Neisseria meningitidis* MC58	conserved hypothetical protein	84%	32384-32745
448	*Neisseria meningitidis* MC58	acetyl-CoA carboxylase, carboxyl transferase beta subunit	84%	706587-706924
447	*Neisseria meningitidis* MC58	tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase	88%	1615773-1616307

Table 5 Continued. Homology of *N. sicca* 4320 to *N. meningitidis* MC58 (EcoRI)

450	*Neisseria meningitidis* MC58	aspartyl-tRNA synthetase	82%	486455-486663
452	*Neisseria meningitidis* MC58	conserved hypothetical protein	94%	1384284-1384903
453	*Neisseria meningitidis* MC58	carboxyphosphonoenolpyruvate phosphonomutase, putative	84%	440874-441094
454	*Neisseria meningitidis* MC58	NlhA outer membrane protein	60%	1119449-1119782
458	*Arabidopsis*	Glycine-rich cell wall structure	54%	
459	*Arabidopsis & Xylella fastidiosa*	Glycine-Rich Cell Wall Structure & endo-1,4-beta-glucanase	54%	
461	*Neisseria meningitidis* MC58	conserved hypothetical protein	88%	2262808-2263467
462	*Neisseria meningitidis* MC58	Enolase (2-phosphoglycerate dehydratase)	85%	
463	*Neisseria meningitidis* MC58	phosphoribosylformylglycinamidine synthase (purL)	61%	2102276-2102412
464	*Neisseria meningitidis* MC58	conserved hypothetical protein	96%	328234-328586
465	*Neisseria meningitidis* MC58	conserved hypothetical protein & acetolactate synthase III, small subunit	90%	1637044-1637514
466	*Neisseria meningitidis* MC58	iron-regulated outer membrane protein (frpB)	63%	2093861-2094190
467	*Neisseria denitrificans*	1,4-alpha-glucan branching enzyme	85%	
468	*Neisseria meningitidis* MC58	conserved hypothetical protein	95%	328806-329007
469	*Neisseria meningitidis* MC58	conserved hypothetical protein	81%	1669656-1670003
470	*Neisseria meningitidis* MC58	conserved hypothetical protein	57%	1049635-1049931
472	*Neisseria meningitidis* MC58	conserved hypothetical protein	60%	847906-848081
474	*Neisseria meningitidis* MC58	phosphoenolpyruvate synthase	88%	649925-650459
475	*Haemophilus parainfluenzae*	Type II restriction enzyme HpaII	70%	
476	*Neisseria meningitidis* MC58	aspartyl-tRNA synthetase	90%	486439-486663
477	*Neisseria meningitidis* MC58	conserved hypothetical protein	57%	
478	*Neisseria meningitidis* MC58	conserved hypothetical protein	57%	1698151-1698313
480	*Neisseria meningitidis* MC58	transporter, putative	93%	1223722-1224165
483	*Neisseria meningitidis* MC58	N-acetylglutamate synthase (argA)	57%	1980070-1980553
486	*Neisseria meningitidis* MC58	malate:quinone oxidoreductase	87%	2215176-2215669
487	*Neisseria meningitidis* MC58	phosphoenolpyruvate carboxylase	89%	2182257-2182895
488	*Neisseria meningitidis* MC58	cell division protein FtsW	86%	431021-431587
		UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide)-pyrophosphoryl-undecaprenol N-acetylglucosamine transferase (murG)	77%	
489	*Neisseria meningitidis* MC58	aminotransferase, class I	83%	1523427-1523748
492	*Neisseria meningitidis* MC58	hypothetical protein	54%	2193816-2194036

Table 5 Continued. Homology of *N. sicca* 4320 to *N. meningitidis* MC58 (EcoRI)

491	*Neisseria meningitidis* MC58	PglB2 (Pilin Glycosylation)	88%	
495	*Neisseria meningitidis* MC58	nitric oxide reductase	88%	1685155-1685446
496	*Neisseria meningitidis* MC58	conserved hypothetical protein	95%	328814-329027
497	*Neisseria meningitidis* MC58	malate:quinone oxidoreductase	88%	2215390-2216216
499	*Neisseria meningitidis* MC58	H.8 outer membrane protein precursor	84%	1583989-1584741
501	*Neisseria meningitidis* MC58	aspartyl-tRNA synthetase	88%	486420-486655
502	*Neisseria meningitidis* MC58	DNA-directed RNA polymerase, beta' subunit	92%	143067-143853
503	*Neisseria meningitidis* MC58	ABC transporter, ATP-binding protein	85%	620181-620890
504	*Neisseria meningitidis* MC58	excinuclease ABC, subunit A	95%	978761-979063
505	*Neisseria meningitidis* MC58	conserved hypothetical protein	89%	2054809-2054980
272	*Neisseria meningitidis* MC58	conserved hypothetical protein	78%	1324230-1324580
203	*Neisseria meningitidis* MC58	aconitate hydratase 2 (acnB)	92%	1633062-1633980
239	*Neisseria meningitidis* MC58	UDP-N-acetylmuramate:L-alanyl-gamma-D-glutamyl-meso-diaminopimelate	91%	1151293-1151914
286	*Neisseria meningitidis* MC58	thiamine biosynthesis protein ThiC (thiC)	92%	996511-998802
207	*Neisseria meningitidis* MC58	phage integrase	60%	
196	*Neisseria meningitidis* MC58	UDP-3-O-3-hydroxymyristoyl N-acetylglucosamine deacetylase (envA)	84%	15280-15439
194	*Neisseria meningitidis* MC58	hemagglutinin/hemolysin-related protein	76%	510783-511114
187	*Neisseria denitrificans*	1,4-alpha-glucan branching enzyme (glgB)	79%	
197	*Neisseria meningitidis* MC58	dimethyladenosine transferase (ksgA)	62%	723237-723356
224	*Neisseria meningitidis* MC58	quinolinate synthetase A (nadA)	92%	398582-398873
226	*Neisseria denitrificans*	1,4-alpha-glucan branching enzyme (glgB)	79%	
253	*Neisseria meningitidis* MC58	leucyl/phenylalanyl-tRNA--protein transferase (aat)	92%	2185598-2185776
193	*Neisseria meningitidis* MC58	imidazoleglycerol-phosphate dehydratase (hisB)	96%	274020-274381
227	*Neisseria meningitidis* MC58	conserved hypothetical protein	94%	328808-329027
265	*Neisseria meningitidis* MC58	conserved hypothetical protein	92%	1773707-1774342
266	*Neisseria meningitidis* MC58	ATP-dependent RNA helicase HrpA, truncation	84%	2126994-2127315
273	*Neisseria meningitidis* MC58	thiamine biosynthesis protein ThiC (thiC)	89%	2159386-2159902
281	*Neisseria meningitidis* MC58	30S ribosomal protein S1 (rpsA)	92%	1319085-1319991
344	*Neisseria meningitidis* MC58	ATP-dependent RNA helicase HrpA, truncation	90%	2127314-2128166
350	*Neisseria meningitidis* MC58	hypothetical protein	60%	628100-628544
353	*Neisseria meningitidis* MC58	PilM protein (Pilus Assembly)	70%	

Table 5 Continued. Homology of *N. sicca* 4320 to *N. meningitidis* MC58 (EcoRI)

354	*Neisseria meningitidis* MC58	malate:quinone oxidoreductase (yojH)	87%	2215360-2216216
355	*Neisseria meningitidis* MC58	fructose-bisphosphate aldolase (cbbA)	85%	1974419-1975019
361	*Neisseria meningitidis* MC58	glucose-6-phosphate isomerase (pgi-1)	72%	343880-344161
365	*Neisseria meningitidis* MC58	hemagglutinin/hemolysin-related protein	79%	510783-511033
368	*Neisseria meningitidis* MC58	lacto-N-neotetraose biosynthesis glycosyl transferase LgtB (lgtB)	59%	2026280-2026501
370	*Neisseria meningitidis* MC58	transposase, IS30 family	80%	2259294-2259780
372	*Neisseria meningitidis* MC58	NADH dehydrogenase I, D subunit (nuoD)	57%	248337-248536
374	*Neisseria meningitidis* MC58	pyruvate kinase II (pykA)	76%	101490-101939
375	*Neisseria meningitidis* MC58	biopolymer transport protein ExbD (exbD)	76%	1811641-1812098
376	*Neisseria meningitidis* MC58	amidophosphoribosyltransferase (purF)	91%	716853-717684
379	*Neisseria meningitidis* MC58	acyl-(acyl-carrier-protein)--UDP-N-acetylglucosamine O-acyltransferase (lpxA)	89%	175430-175936
380	*Neisseria meningitidis* MC58	elongation factor G (EF-G) (fusA)	97%	148638-148830
382	*Neisseria meningitidis* MC58	conserved hypothetical protein	63%	826822-826928
383	*Neisseria meningitidis* MC58	pilQ protein (pilQ)	62%	1902283-1902691
384	*Neisseria meningitidis* MC58	conserved hypothetical protein	82%	631247-631935
385	*Neisseria meningitidis* MC58	histidinol dehydrogenase (hisD)	81%	1642628-1643498
386	*Neisseria meningitidis* MC58	ribonuclease II-related protein	93%	289238-290119
388	*Neisseria meningitidis* MC58	adenylosuccinate lyase (purB)	94%	291866-292033
389	*Neisseria meningitidis* MC58	membrane protein	75%	619359-620118
397	*Neisseria meningitidis* MC58	cytochrome c-type biogenesis protein, putative	70%	1893833-1894401
392	*Neisseria meningitidis* MC58	membrane protein	75%	619359-620118
326	*Neisseria meningitidis* MC58	phosphoribosylformylglycinamidine synthase (purL)	86%	2101929-2102438
300	*Neisseria meningitidis* MC58	phosphoenolpyruvate synthase (ppsA)	87%	649924-650749
301	*Neisseria meningitidis* MC58	thiamine biosynthesis protein ThiC (thiC)	80%	2161405-2161741
306	*Neisseria meningitidis* MC58	ATP synthase F0, C subunit (atpE)	87%	2036783-2037630
312	*Neisseria meningitidis* MC58	lysyl-tRNA synthetase, heat inducible (lysU)	84%	1461330-1461713
340	*Neisseria meningitidis* MC58	transcription-repair coupling factor (mfd)	92%	1298027-1298753
315	*Neisseria meningitidis* MC58	argininosuccinate synthase (argG)	73%	2237078-2237195
399	*Neisseria meningitidis* MC58	xanthine/uracil permease family protein	54%	945511-946012
401	*Neisseria meningitidis* MC58	acyl-(acyl-carrier-protein)--UDP-N-acetylglucosamine O-acyltransferase (lpxA)	88%	175431-175936

Table 6. Homology of *N. sicca* 4320 to *N. meningitidis* MC58 (AccI)

Clone	Homology To Strain	Protein Function	%-N	MC58 Pos.
A1	*Neisseria meningitidis* MC58	16S RNA methyltransferase (rsmB)	86%	119586 - 120045
A2	*Neisseria meningitidis* MC58	malate:quinone oxidoreductase (yojH)	89%	2215724 - 2216049
A4	*Neisseria meningitidis* MC58	hemolysin activation protein HecB, putative	89%	1841629 - 1842375
A5	*Neisseria meningitidis* MC58	DNA polymerase III, subunits gamma and tau, programmed frameshift (dnaX)	86%	1484583 - 1484973
A6	*Neisseria meningitidis* MC58	phytoene synthase, putative	58%	1140241 - 1140449
A10	*Neisseria meningitidis* MC58	cytochrome c-type biogenesis protein, putative	68%	1712209 - 1712370
A12	*Acinetobacter sp.*	4-oxalocrotonate decarboxylase	81%	
A13	*Neisseria meningitidis* MC58	leucyl-tRNA synthetase (leuS)	87%	1996955 - 1997531
A14	*Neisseria meningitidis* MC58	phosphoribosylaminoimidazole-succinocarboxamide synthase (purC)	81%	785065 - 785516
A15	*Neisseria meningitidis* MC58	fructose-bisphosphate aldolase (cbbA)	84%	377588 - 377704
A16	*Neisseria meningitidis* MC58	DNA polymerase holoenzyme chi subunit, putative	63%	1627668 - 1627805
A17	*Neisseria meningitidis* MC58	acetyl-CoA carboxylase, biotin carboxylase (accC)	58%	179821 - 179889
A19	*Neisseria meningitidis* MC58	leucyl-tRNA synthetase (leuS)	86%	1997747 - 1998111
A20	*Neisseria meningitidis* MC58	acetyl-CoA carboxylase, biotin carboxylase (accC)	58%	179820 - 179889
A21	*Neisseria meningitidis* MC58	sulfite reductase hemoprotein, beta-component (cysI-1)	72%	1909702 - 1909732
A22	*Neisseria meningitidis* MC58	fructose-bisphosphate aldolase (cbbA)	86%	1974488 - 1975193
A24	*Neisseria meningitidis* MC58	transposase, IS30 family	92%	234022 -234613
A25	*Neisseria meningitidis* MC58	leucyl-tRNA synthetase (leuS)	93%	1099218 - 1099323
A26	*Neisseria meningitidis* MC58	TonB-dependent receptor	69%	1925812 - 1926336
A27	*Neisseria meningitidis* MC58	lipopolysaccharide glycosyl transferase (lgtG)	56%	2150150 - 2150466
A28	*Neisseria meningitidis* MC58	organic solvent tolerance protein, putative	56%	283498 - 283749
A29	*Neisseria meningitidis* MC58	ATP-dependent Clp protease, ATP-binding subunit ClpA (clpA)	65%	862620 - 862687
A103	*Neisseria meningitidis* MC58	trans-sulfuration enzyme family protein	89%	1672254 - 1672906
A105	*Neisseria meningitidis* MC58	DNA polymerase holoenzyme chi subunit, putative	62%	1627668 - 1627826
A106	*Neisseria meningitidis* MC58	zinc uptake regulation protein, putative (zur)	70%	1276533 - 1276699
A107	*Neisseria meningitidis* MC58	NosR-related protein	69%	605345 - 605786
A109	*Neisseria meningitidis* MC58	alanyl-tRNA synthetase (alaS)	54%	1657321 - 1658026
A111	*Neisseria meningitidis* MC58	DNA polymerase III, subunits gamma and tau, programmed frameshift (dnaX)	88%	1484494 - 1484973
A114	*Neisseria meningitidis* MC58	16S RNA methyltransferase (rsmB)	85%	119586 - 120066
A115	*Neisseria meningitidis* MC58	hypothetical protein	59%	1237043 - 1237185

Table 6 Continued. Homology of *N. sicca* 4320 to *N. meningitidis* MC58 (AccI)

A116	*Neisseria meningitidis* MC58	DNA polymerase holoenzyme chi subunit, putative	61%	1627668 - 1627799
A30	*Neisseria meningitidis* MC58	sodium- and chloride-dependent transporter	66%	1783659 - 1784095
A31	*Neisseria meningitidis* MC58	leucyl-tRNA synthetase (leuS)	91%	1996147 - 1996685
A32	*Neisseria meningitidis* MC58	transposase, IS30 family	92%	1854573 - 1855147
A33	*Neisseria meningitidis* MC58	ADP-heptose synthase, putative	61%	847926 - 848081
A34	*Neisseria meningitidis* MC58	acetyl-CoA carboxylase, biotin carboxylase (accC)	57%	1965096 - 1965409
A35	*Neisseria meningitidis* MC58	transposase, IS30 family	70%	926278 - 926616
A36	*Neisseria meningitidis* MC58	ATP-dependent Clp protease, ATP-binding subunit ClpA (clpA)	65%	862620 - 862687
A37	*Neisseria meningitidis* MC58	serine-type peptidase	63%	2110268 - 2110445
A38	*Neisseria meningitidis* MC58	ATP-dependent Clp protease, ATP-binding subunit ClpA (clpA)	65%	862620 - 862687
A39	*Neisseria meningitidis* MC58	zinc uptake regulation protein, putative (zur)	75%	1276265 - 1276699
A41	*Neisseria meningitidis* MC58	diaminopimelate decarboxylase (lysA)	59%	820001 - 820087
A44	*Neisseria meningitidis* MC58	zinc uptake regulation protein, putative (zur)	85%	1276215 - 1276699
A47	*Neisseria meningitidis* MC58	conserved hypothetical protein	79%	450054 - 450352
A48	*Neisseria meningitidis* MC58	fructose-bisphosphate aldolase (cbbA)	89%	1974039 - 1974505
A49	*Neisseria meningitidis* MC58	conserved hypothetical protein	71%	2253054 - 2253563
A50	*Neisseria meningitidis* MC58	fructose-bisphosphate aldolase (cbbA)	89%	1974039 - 1974509
A51	*Neisseria meningitidis* MC58	macrophage infectivity potentiator	65%	1627126 - 1627208
A56	*Neisseria meningitidis* MC58	NosR-related protein	79%	605345 - 605785
A57	*Neisseria meningitidis* MC58	acetyl-CoA carboxylase, biotin carboxylase (accC)	58%	1965096 - 1965402

Table 6 Continued. Homology of *N. sicca* 4320 to *N. meningitidis* MC58 (AccI)

A59	*Neisseria meningitidis* MC58	zinc uptake regulation protein, putative (zur)	84%	1276166 - 1276699
A62	*Neisseria meningitidis* MC58	acetyl-CoA carboxylase, biotin carboxylase (accC)	58%	1965096 - 1965402
A63	*Neisseria meningitidis* MC58	oligopeptidase A (prlC)	89%	218655 - 219256
A64	*Neisseria meningitidis* MC58	DNA polymerase holoenzyme chi subunit, putative	62%	1627668 - 1627841
A66	*Neisseria meningitidis* MC58	conserved hypothetical protein	75%	1532648 - 1532752
A67	*Neisseria meningitidis* MC58	CTP synthase (pyrG)	87%	1613777 - 1613988
A69	*Neisseria meningitidis* MC58	cytochrome c-type biogenesis protein, putative	73%	1891384 - 1892260
A70	*Neisseria meningitidis* MC58	arginyl-tRNA synthetase (argS)	93%	1555739 - 1555951
A72	*Neisseria meningitidis* MC58	trans-sulfuration enzyme family protein	92%	1672069 - 1672906
A73	*Neisseria meningitidis* MC58	lipopolysaccharide glycosyl transferase (lgtG)	56%	2150207 - 2150465
A74	*Neisseria meningitidis* MC58	acetyl-CoA carboxylase, biotin carboxylase (accC)	58%	1965096 - 1965402
A84	*Neisseria meningitidis* MC58	hemagglutinin/hemolysin-related protein	67%	1219938 - 1220297
A85	*Neisseria meningitidis* MC58	IS1106 transposase	90%	1004255 - 1004459
A95	*Neisseria meningitidis* MC58	conserved hypothetical protein	72%	9634 - 10330
A98	*Neisseria meningitidis* MC58	hypothetical protein	60%	2098592 - 2098736
A99	*Neisseria meningitidis* MC58	guanosine-3`,5`-bis(diphosphate) 3`-pyrophosphohydrolase (spoT)	80%	1733316 - 1733980

Overlay of EcoRI and AccI fragments on *N. meningitidis* MC58 chromosomal map.

The high levels of homology of *N. sicca* 4320 DNA sequence to the *N. meningitidis* MC58 chromosome allowed for further genomic comparison. Since the MC58 chromosome has been sequenced and annotated, it was used as a backbone to organize homologous 4320 regions. The overlay of 4320 sequences onto the MC58 chromosome allows for comparison of the two chromosomes as well as observing if the cloning process was effective in sampling random regions of the 4320 chromosome.

Each *N. sicca* 4320 EcoRI and AccI chromosomal fragment was used as query sequence to search the *N. meninigtidis* MC58 database at TIGR. The search results yielded the position of the homologous gene on the meningococcus chromosome. The locations of all homologs were complied using the program Vector NTI.

These resulting 4320 chromosomal maps for the EcoRI and AccI fragments are respectively shown in Figures 3 and 4. Each clone containing the 4320 chromosomal fragments is represented by a mark on the map. The results indicate that the clones contained random sequences from all regions of the chromosome, as the clone marks extend around the map. Closer examination of the maps reveals that regions exist to which multiple clones mapped as well as regions lacking the presence of any clones. For example, overlay of both restriction maps shows several holes located between clones 443 and 374, between 270 and 452, and between 452 and 312. These holes could either represent areas of unique sequence to 4320, or areas containing genes important for pathogenesis.

Although there is homology of 4320 sequences to those of MC58 and these sequences are randomly distributed on the circumference of the MC58 chromsome it can

not be concluded that these chromosomes have the same organization. To show that the organizations of the chromosomes are similar, linkage analysis needs to be performed using clones that are adjacent on the map. For example, PCR should be utilized to show that clone 161 is located near clone 383.

Linkage analysis also should be performed to determine if the 4320 clones 368, A27, and A73 are linked, or if they are at the same loci as shown in Fig 3 and 4. These clones show low homologies to neisserial glycosyl transferases. Linkage of these *lgtB* and *lgtG* homologs in the 4320 chromosome would be indicative of diverse LOS biosynthetic genes and organizations in the commensal *Neisseria*.

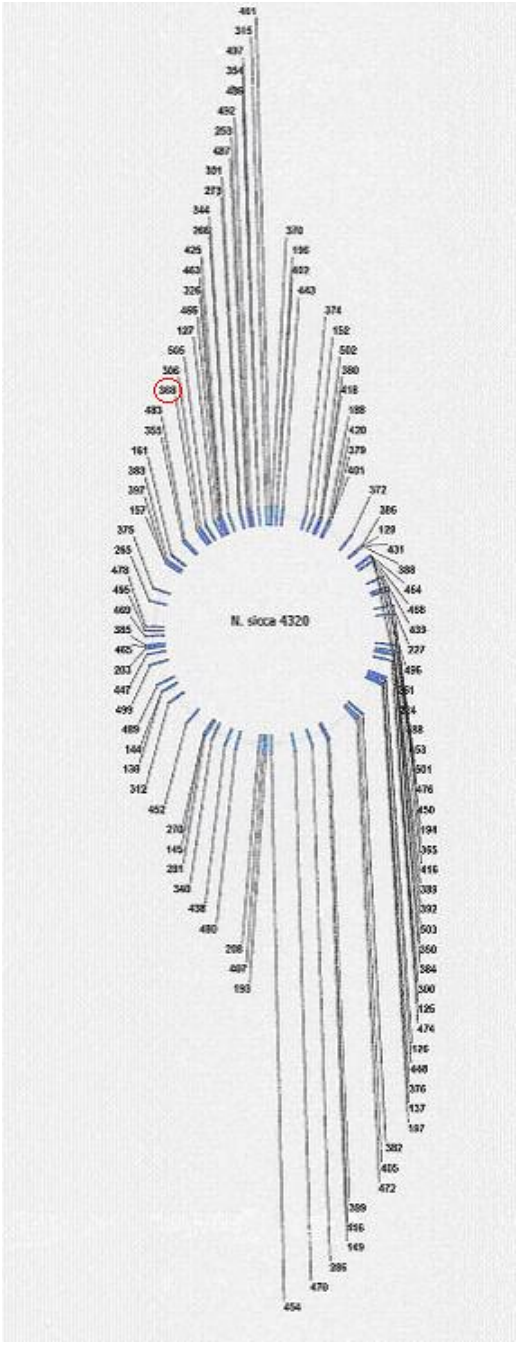


Figure 3. Chromosomal map of *N. sicca* 4320 generated from EcoRI shotgun clone sequences. Each number corresponds to a clone containing a random insertion of 4320 chromosomal DNA into the EcoRI site of pUC19. Encircled is a clone containing a potential *lgtB* homolog.

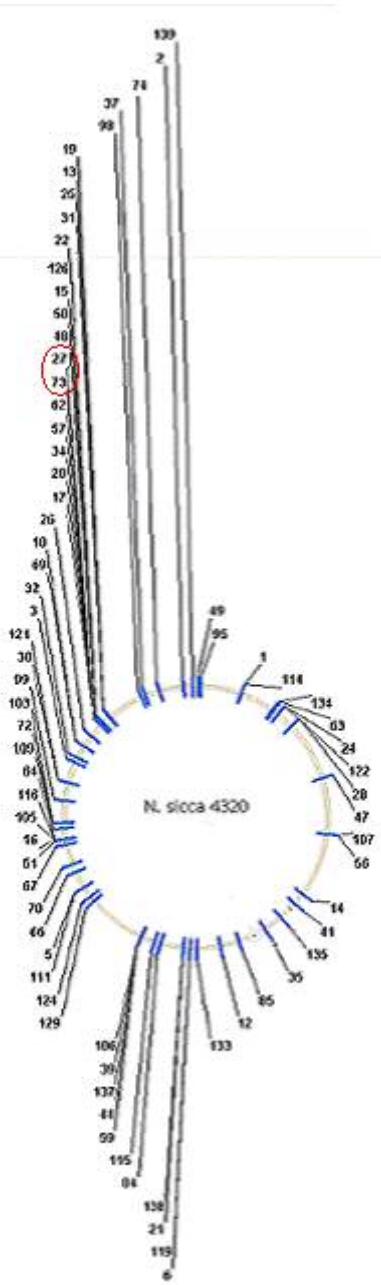


Figure 4. Chromosomal map of *N. sicca* 4320 generated from AccI shotgun clone sequences. Each number corresponds to a clone containing a random insertion of 4320 chromosomal DNA into the AccI site of pUC19. Encircled are clones containing potential *lgtG* homologs.

Analysis of the presence of *lpt3* and *lpt6* in *N. sicca* strains.

Neisserial LOS can be decorated with phophoethanolamine (PEA) residues, modifications that can alter the organism's pathogenic potential. A study showed that *N. meningitidis* expressing LOS with PEA attached to the 6-position of HepII serves as a receptor for complement component C4b. As a result, neisserial strains expressing LOS modified with this PEA modification are more serum sensitve then strains expressing PEA modified LOS at 6-HepII (102). Because PEA decorations contribute to neisserial pathogenicity, the presence of transferase genes mediating these additions was analyzed in *N. sicca* by Southern hybridization analysis.

The genes *lpt3* and *lpt6* have been suggested to encode PEA transferases that mediate PEA decoration respectively at the 3 and 6 positions of HepII (81) (139). Biochemical evidence that Lpt3 actually functions as a PEA transferase is provided in this dissertation. The DNA containing these genes were amplified using primers Lpt-6-F and Lpt-6-R for *lpt6* and PTE-4 and PTE-5 for *lpt3*. PCR products were labeled with DIG and used as probes to detect the presence of these genes in *N. sicca* strains as well as the gonococcus and meningococcus.

The results of the hybridization studies are shown in Figs. 5 and 6. Southern analysis to detect *lpt3* revealed strong hybridization signals in the gonococcal and meningococcal chromosomes. A variable *lpt3* presence was seen in the *N. sicca* chromosomes, as *N. sicca* and *N. sicca* 4318 chromosomes failed to hybridize with *lpt3*. In addition, the *lpt3* probe hybridized to fragments of different *EcoRI* sizes, indicating genetic variation within the *lpt3*. Significant variation is also seen in the presence of *lpt6* in the examined *Neisseria*. Hybridization signals were seen across the gonococcal and

meningococcal strains. *N. sicca* strains also showed hybridization with the *lpt6* probe. However, several different restriction patterns were seen within the *N. sicca* strains, exhibiting genetic diversity within *lpt6*. Additionally, no *lpt6* hybridization signal was detected for *N. sicca* 4319 and *N. sicca* 4320, strains that were isolated from fatal endocarditis patients and synthesize LPS. These results indicate that significant differences exist within the PEA transferase genes across the *Neisseria* that might account for varying pathogenicities.

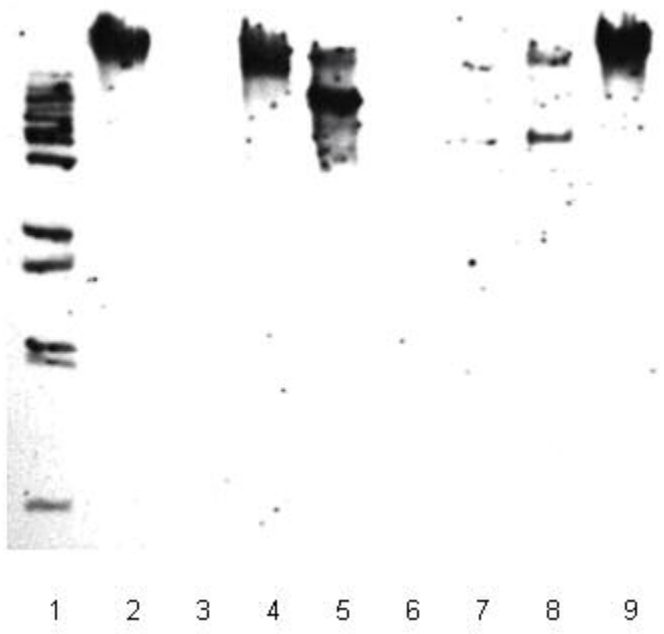


Figure 5. Southern hybridization to test for the presence of *lpt3* in *Neisseria* strains. Lane 1 shows the DIG labeled λ BstEII digested standard. Lanes 2 through 9 show EcoRI digested chromosomes including *N. gonorrhoeae* F62, *N. sicca*, *N. sicca* 19, *N. sicca* 342, *N. sicca* 4318, *N. sicca* 4319, *N. sicca* 4320, and *N. meningitidis* 53418. The membrane was probed with DIG labeled *N. gonorrhoeae* F62 *lpt3*.

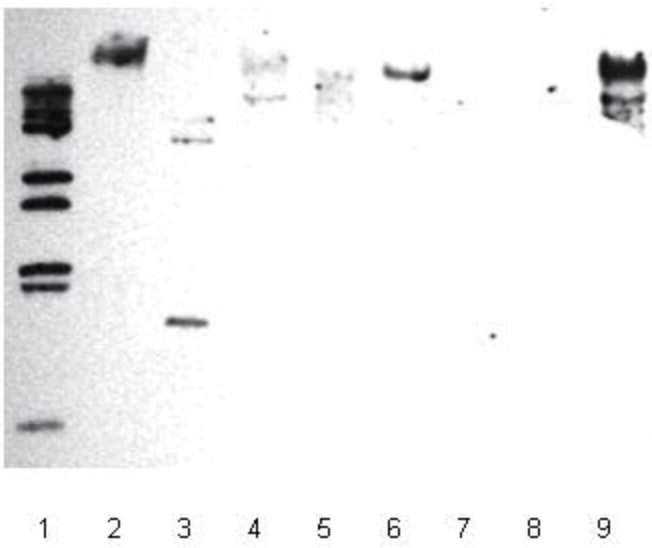


Figure 6. Southern hybridization to test for the presence of *lpt6* in *Neisseria* strains. Lane 1 shows the DIG labeled λ BstEII digested standard. Lanes 2 through 9 show EcoRI digested chromosomes including *N. gonorrhoeae* F62, *N. sicca*, *N. sicca* 19, *N. sicca* 342, *N. sicca* 4318, *N. sicca* 4319, *N. sicca* 4320, and *N. meningitidis* 53418. The membrane was probed with DIG labeled *N. gonorrhoeae* F62 *lpt6*.

DISCUSSION

In this study, a bioinformatic analysis of the *N. sicca* 4320 chromosome was performed to analyze the genomic content of this strain. Sequence comparison of these two neisserial classes on a whole chromosomal level has yet to be reported. For this analysis, random fragments from the 4320 chromosome were sequenced and aligned with the *N. meningitidis* MC58 chromosome.

The sequence analysis results allow for several conclusions to be made. First, *N. sicca* 4320 is a member of the *Neisseria* genus because its chromosomal fragments showed significant homologies with *N. meningitidis* MC58 DNA. DNA sequence analysis is commonly used for the identification and detection of *Neisseria*. To diagnose neisserial infections, genus specific DNA sequences are used to create primers to PCR amplify neisserial DNA (77). Since, 4320 nucelotide sequences showed significant identity to MC58 sequences, this strain belongs to the *Neisseria* genus.

The results of the sequence analysis also reveal that there is not a significant difference in the chromosomal contents of the commensal and pathogenic *Neisseria*. *N. sicca* 4320 sequences were overlayed on the *N. meningitidis* MC58 chromsomal map. Although the map is largely incomplete, as only 3% of the estimated chromosome was sequenced, it showed a distribution of sequences throughout the whole chromosome. In addition, few fragments were found that were not homologous to MC58. Therefore, these two strains are likely genetically similar. Although a linkage analysis of adjacent clones was not performed, chromosomal organizational similarity between these two strains is not unexpected. Both *N. sicca* and *N. meningitidis* are frequent colonists are the oropharyx. Because these species share the same niche and are naturally competent,

genetic information can be transferred between these organisms. Over time a clonal population structure would be expected (118).

Additionally, because *N. meningitidis* and *N. sicca* colonize the same niche, these organisms are predicted to encode similar surface proteins that mediate adherence to host cells. Sequence analysis of 4320 fragments indicate that it contains the genetic potential to produce surface molecules that interact with the host including LOS, pilin, iron binding proteins, and H8. This result is consistent with previous studies of other commensal *Neisseria*. This group of organisms is known to express molecules, LOS, Opa, Por, iron binding proteins, pilin and Rmp, that are similar to those produced by the pathogenic strains (41,80,126,129,141).

This study also suggests that there could be a genetic history shared by *N. sicca* 4320 and members of the *Haemophilus* species, which inhabit the same niche. 4320 chromosomal DNA contains a region that is homolgous to the HpaII restriction enzyme of *Haemophilus parainfluenzae*. Other studies have also found evidence of genetic exchange between the *Neisseria* and *Haemophilus influenzae*. The meningococcus obtained the superoxide dismutase encoding virulence gene *sodC* from *H. influenzae*, as indicated by the presence of the 9 bp *Haemophilus* uptake sequence found upstream of the acquired gene. The neisserial uptake sequence has also been found in *Haemophilus*. For example, Brazilian Purpuric Fever is caused by *H. influenzae*. Infection with this specific strain results in septicaemia, much like a meningococcal infection (76). It is believed that the meningococcus likely donated a gene to *H. influenzae* that encodes an epithelial cell invasion protein (79). Exchange between the *Neisseria* and *Haemophilus* is also evident in the presence of phosphorylcholine on commensal LOS rather than on the

pathogenic pili. *licABCD* was found only in commensal strains and is inserted between two genes that are adjacent in pathogenic *Neisseria*. Each gene in the locus has significant homology to those of *H. influenzae* (115). These examples demonstrate that *Neisseria* can acquire new genes from outside of the genus, and that these genes can have access to the global *Neisseria* pool.

It is also interesting to note that like the *Neisseria*, it has been reported that *Haemophilus* strains produce LOS. However, it has recently been reported that *H. ducreyi* is able to express LPS containing a polylactosamine repeat (110). Also, it has been found that under certain growth conditions other *Haemophilus* strains can synthesize LOS containing a single O-antigen repeat consisting of four sugar residues. As in the synthesis of LPS in other gram-negative species, this addition is mediated by an isoprenoid carrier that links the O-repeat to the LOS core. (59). Although the genes necessary for the linkage of the O-repeat to this carrier, *wecA* or *wbaP,* were not found along with other genes required for O-repeat addition to the LOS, *wzx* (flippase), *wzy* (polymerase), and *wzm*/*wzt* (ABC transporter), they still may be present in the 4320 chromosome. The fragments analyzed in this study only represented a fraction of the genetic capacity of the strain, 3%. Approximately 6000 more clones containing random sequence fragments are required to cover the 4320 chromosome. Since the *Haemophilus* genus contains genes that are required for the addition of the O-repeat to LOS and 4320 is known to have acquired *Haemophilus* DNA, hybridization studies could be performed to test for the presence of these genes in 4320.

Prior to this study, it was known that *N. sicca* 4320 produced LPS, as exhibited by the presence of an O-antigen repeat upon SDS-PAGE analysis (108). In addition, this

study found differences in the synthesis of LOS upon comparison to the pathogenic *Neisseria*. For example, sequence analysis of two 4320 chromosomal fragments showed low homology to *lgtB* and *lgtG* of *N. meningitidis* MC58. This low level of homology is consistent with other studies that have either not detected LOS biosynthesis genes in commensal strains, or have shown low homology levels (6,141). To determine if these genes play a role in the synthesis of 4320 LPS, mutations of the genes need to be constructed in the chromosome. The procedure to genetically manipulate 4320 is described in this work and will allow for the *lgtB* and *lgtG* homologs to be studied.

This study has also shown genetic variability within the *lpt3* and *lpt6* phosphoethanolamine transferase genes. Both genes are present in the gonococcal and meningococcal chromosomes studied, but have a scattered presence within *N. sicca* strains. It is noticeable that both *N. sicca* 4319 and 4320 are responsible for fatal cases of endocarditis and express LPS. Stringent Southern hybridization analysis, requiring homology at an 85% level, failed to detect *lpt6* in these isolates, but detected *lpt3*. PEA decoration of LOS at 6-HepII of the meningcoccus serves as a receptor for complement component C4b, conferring complement sensitivity to the *Neisseria* with this modification (102). Because both 4319 and 4320 were complement resistant and able to disseminate in the bloodstream, the lack of PEA modification at 6-HepII, the potential expression of PEA at 3-HepII, and the production of LPS likely mediate the survival of these isolates in the bloodstream, therefore altering their pathogenicity.

CHAPTER 4. STRUCTURAL DETERMINATION OF *N. SICCA* 4320 LIPOPOLYSACCHARIDE

INTRODUCTION

A previous study of LOS biosynthesis in the *Neisseriaceae* showed diversity in this class of surface antigen. Some commensal strains produced truncated versions of LOS while others produced longer polymers forming an O-repeat antigen (108). It was hypothesized that *N. sicca* 4320 actually produced LPS in which an LOS core is present and to which an O-repeat antigen is added. The goal of this chapter is to determine the structure of the LPS produced.

A few studies have shown that both the *Neisseria* and *Haemophilus* have the ability to modify their LOS molecules with polylactosamine, and these strains seem to possess increased virulence. For example, when *N. gonorrhoeae* MS11mkC was used to inoculate healthy male volunteers, 43% of the volunteers developed urethritis, compared to the average rate of 22%. In the same set of experiments it was found that sialyated MS11mkC were unable to infect the urethra. Fresh isolates taken from the male volunteers who had contracted urethritis produced LOS molecules with increased mass compared to the original inoculum which expressed paraglobsyl and gangliosyl LOS (111,114). MALDI-TOF and ES-MS analyses of these molecules showed that polylactosamine repeats had been added to the LOS (64). It was speculated that the addition of the repeats contributed to the increase in virulence of the MS11 strain. Results analgous to the previously described set of experiments were generated while studying *Haemophilus ducreyi*, a causative agent of genital ulcers. It was determined that a linear polymer consisting of polylactosamine was added to the LOS molecule (36,110).

The addition of polylactosamine to the LOS of *Neisseria* and *Haemophilus* is significant because it is associated with an increase in pathogenicity. The ability of these organisms to synthesize the addition has only recently been discovered. The addition of polylactosamine to the gonococcal LOS was seen on initial *in vivo* subculture, as the decoration is lost after extended *in vitro* passage. However, because of their role in increasing virulence, study of these types of molecules as well as the genetic mechanism underlying their synthesis is important, especially in the design of an LOS based vaccine. Like the gonococcus, the synthesis of LPS by *N. sicca* 4320 likely contributed to its increased virulence. However, unlike the gonococcus, 4320 stably expresses LPS *in vitro*. It is thought that this isolate may possess the same genetic mechanism as the gonococcus to synthesize its LPS, and therefore can be used to study LPS biosynthesis in the pathogenic *Neisseria*.

RESULTS

MALDI-TOF analysis showing *N. sicca* 4320 produces two separate molecules – LOS and LPS.

LPS was extracted and purified from *N. sicca* 4320 cells according to the hot-phenol extraction method (61). Shown in Fig 7 is the purified molecule that has been electrophoresed on an SDS-PAGE gel. Evident on this gel is the repeating pattern that is characteristic of LPS. To analyze the purified molecule, MALDI-TOF mass spectrometry was performed in by the Griffiss laboratory (VA Hospital, San Francisco, CA). The resulting spectra is shown in Fig 8. As shown in the spectra, three major peaks were seen. The ion represented by the 951.50 *m/z* is a mass consistent with neisserial lipid A (64). Based on the known components of other neisserial LOS molecules, the 2414.81 *m/z* peak suggests an LOS corresponding with the predicted structure of $(Hex)_2(Hep)_3(PEA)_1(KDO)_1$(lipoidal moiety). The 2222.65 *m/z* ion is consistent with the loss of a Heptose (-192 Da) residue from the LOS. The LOS molecule is shown as the dark prominent band on the SDS-PAGE gel. Fragmentation of the LOS occurred during MALDI-TOF analysis, resulting in the three described peaks. A closer examination of the 4320 LOS shows that this molecule is different from other traditional neisserial LOS molecules. The SDS-PAGE gel in Fig. 9 shows that the 4320 LOS molecule has a comparable mobility to 2426 *m/z N. gonorrhoeae* F62ΔLgtAlpt3::Tn5 LOS. However, although the masses of these molecules are similar, the predicted structures are different, with F62ΔLgtAlpt3::Tn5 LOS composed of $(Hex)_2(Hep)_2(PEA)_1(KDO)_1$(lipoidal moiety). These structures alone demonstrate the diversity of LOS molecules that exist within the *Neisseria* genus.

A closer look at the spectra shows a repeating pattern with successive peaks that are separated by 349 *m/z*. Each peak corresponds to one band composing the ladder pattern on the LOS gel. This repeating pattern is evident at values smaller than the 951 *m/z* lipid A peak. Because the repeating pattern contains molecules with masses less than that of lipid A, two possibilities exist. First, the purification procedure could have cleaved the O-repeat from the LOS, if the O-repeat extends from the LOS on the 4320 cell surface. Second, the repeating pattern may not be anchored to the membrane by traditional *Neisseria* lipid A, and may be a component of a separate molecule. The successive subtraction of 349 mass units from the repeating pattern in the spectra leaves a molecule of either 502 or 153 *m/z*, representing a possible different anchoring mechanism for the O-repeat. Therefore, *N. sicca* 4320 could express two separate molecules on its surface, LOS and LPS.

Figure 7. Isolated and purified *N. sicca* 4320 LPS. *N. sicca* 4320 LPS was electrophoresed on a 16.5% SDS-PAGE gel and silver stained. The prominent high mobility band represents the LOS molecule, while the low mobility molecules represent successive additions of O-antigen.

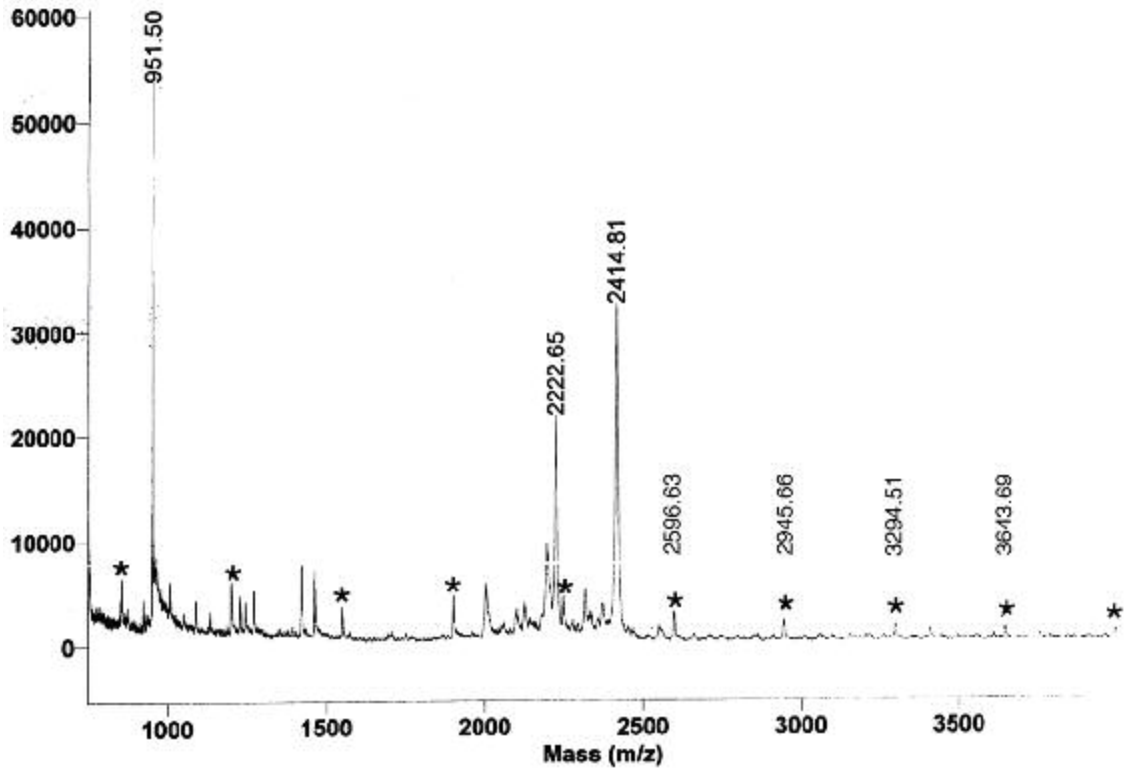


Figure 8. MALDI-TOF analysis showing that *N. sicca* 4320 produces LOS and LPS. The spectra contains the purified O-deacylated LOS and LPS of *N. sicca* 4320. The masses of the abundant fragments are indicated at the top of the corresponding peaks. Starred peaks represent a series of molecular ions separated by a constant 349 *m/z* mass difference. The spectra was generated in the laboratory of Dr. Griffiss at UCSF.

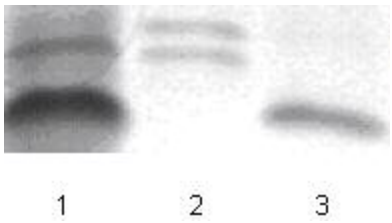


Figure 9. SDS-PAGE comparing *N. sicca* 4320 LOS to previously characterized gonococcal LOS molecules. The gel contains the LOS molecules from strains *N. sicca* 4320, *N. gonorrhoeae* F62, *N. gonorrhoeae* F62ΔLgtAlpt3::Tn5 respectively in lanes 1-3.

MALDI-MS and ESI-MS showing *N. sicca* 4320 LPS contains a disaccharide repeat.

To determine the components of the repeating unit, MALDI-MS and ESI-MS was performed by Vern Reinhold's lab at the University of New Hampshire. The spectra resulting from the MALDI-MS analysis on the 4320 LPS sample is shown in Fig. 10. The sample was treated with acid to remove the lipoidal moiety and methylated, prior to analysis. This profile shows peaks that differ from each other by 419 *m/z* units. The difference is best viewed in the spectra between 2400 and 4800 *m/*z, a region free from the expected size of the OS. In this region peaks corresponding to ions of 2651.1, 3069.9, 3490.1, 3909.2, 4328.6, and 4747.9 *m/z* exist, forming a series of ions that differ by 419 *m/z*. To further establish the components of the 419 *m/z* repeating unit, ESI-MS was performed on pyrazole derivitized PS with the results shown in Fig 11. ESI-MS allows for further molecular fragmentation to provide evidence that the repeat components. This spectrum confirmed the presence of the 419 *m/z* repeating unit along with mass differences that would be consistent with the lack of a pentose and HexNAc. For example, the 1809.8 *m/z* peak corresponds to four repeating units linked to pyrazole. The ion represented by the 1635.8 *m/z* peak is 174 mass units smaller than the 1809.8 *m/z* ion, consistent with the loss of a methlyated pentose. Compared to the 1635.8 *m/z* ion, the 1390 *m/z* ion is 245 mass units smaller, or the mass of methylated HexNAc. The 1216.7 and 971.8 *m/z* ions also show the same repeating 174 and 245 mass differences. The predicted structures of these fragments are shown in Fig 12.

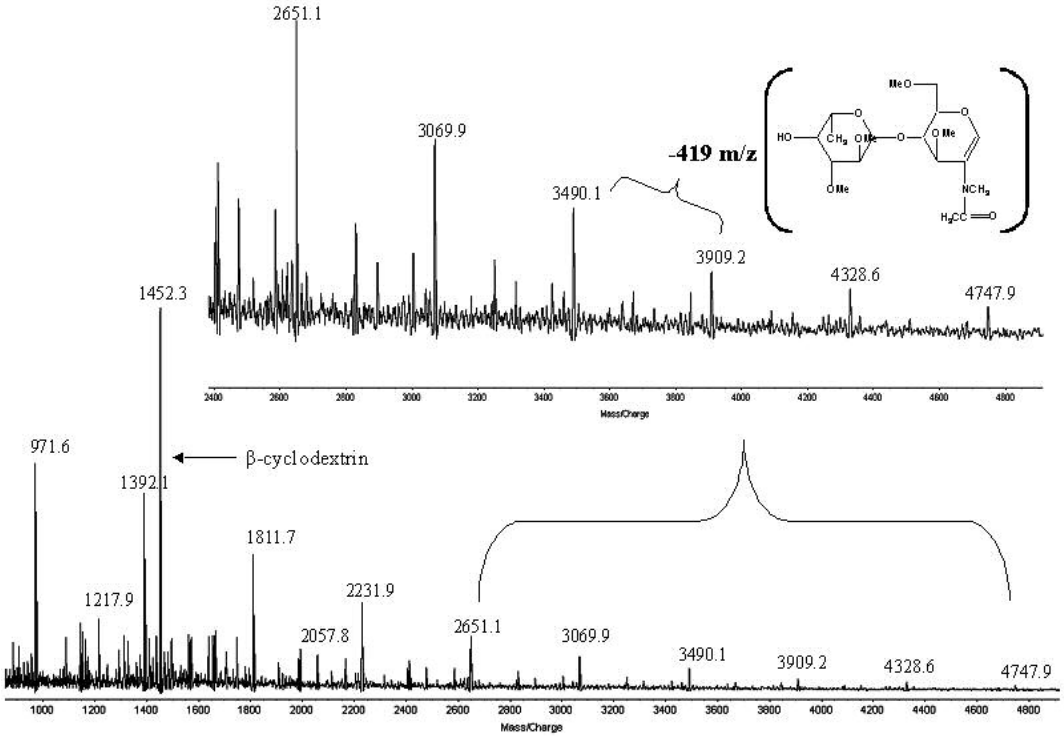


Figure 10. MALDI-TOF analysis showing that *N. sicca* 4320 LPS contains a dissacharide repeat. The spectra shows the profile of *N. sicca* 4320 PS produced by MALDI-MS. The β-cyclodextrin standard is shown and the masses of the abundant fragments are indicated at the top of the corresponding peaks. The region from 2400 to 4800 *m/z* is enlarged to allow for a clear view of the peaks representing the O-repeat ions. The spectra was generated in the laboratory of Dr. Vern Reinhold at the University of New Hampshire.

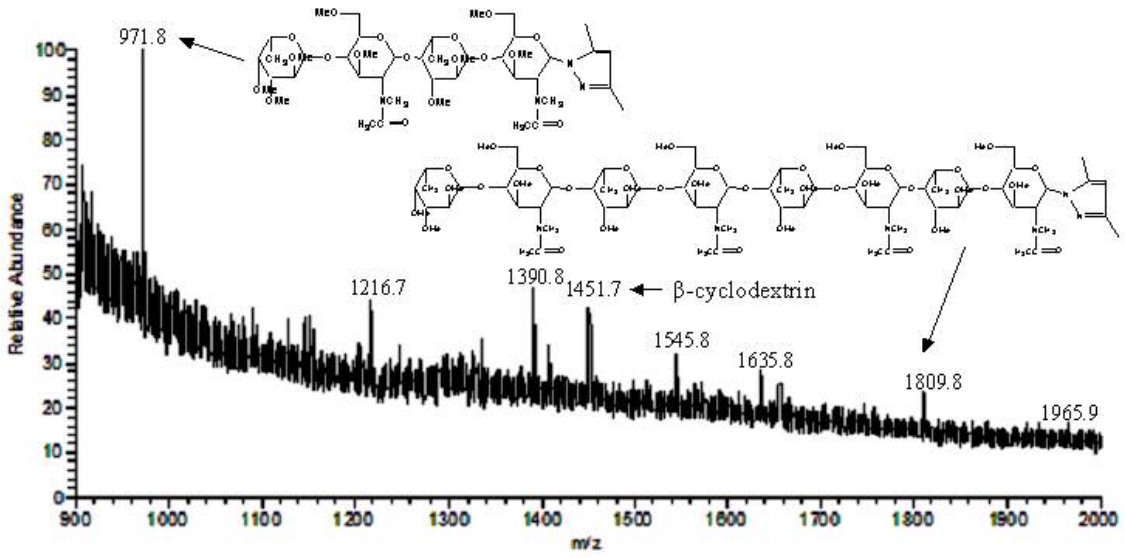


Figure 11. ESI-MS analysis showing a Pentose – HexNAc repeat. The spectra in shows the results of ESI-MS of the pyrazole derivitized methylated PS. The masses of the abundant fragments are displayed over the corresponding peaks. Structures representing the 971.8 and 1809.8 *m/z* ions are also shown. The spectra was generated in laboratory of Dr. Vern Reinhold at the University of New Hampshire.

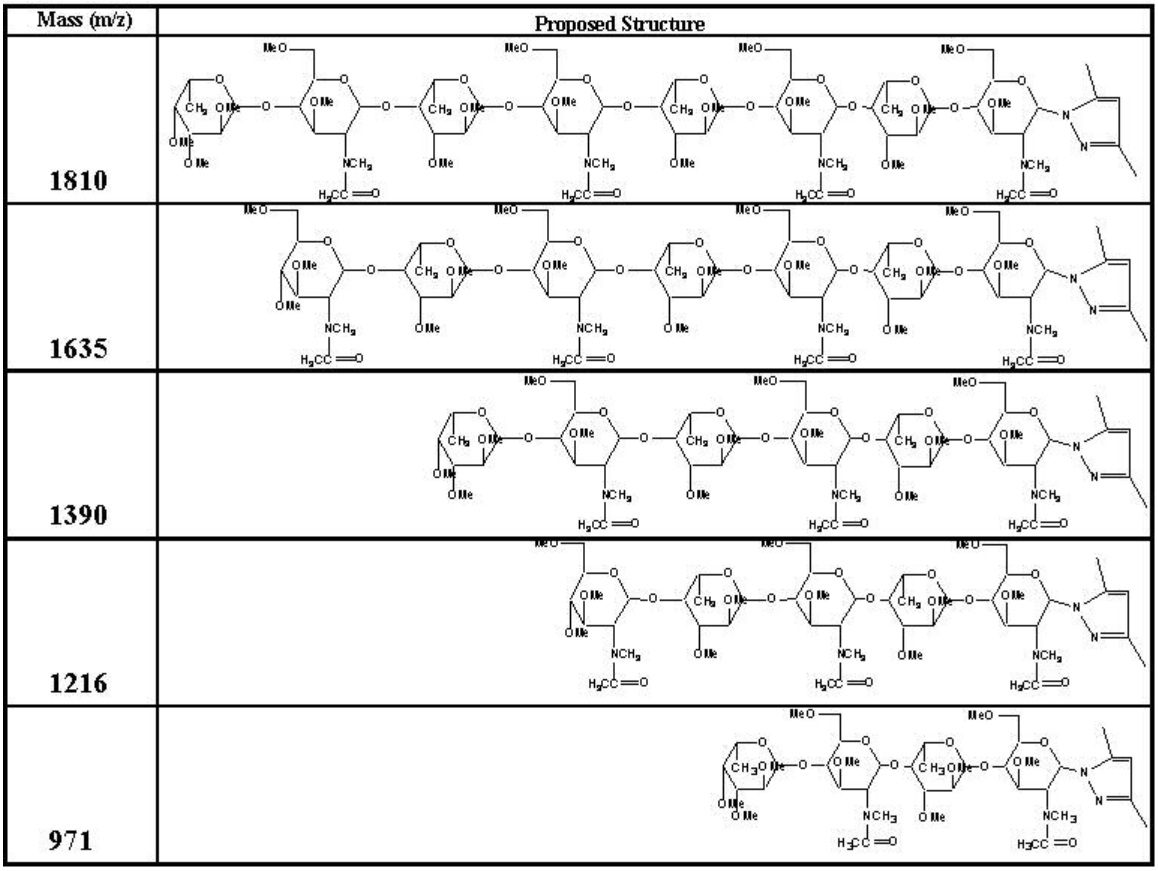


Figure 12. Proposed structures of *N. sicca* 4320 LPS repeat molecules resulting from ESI-MS analysis. The mass units in the right column correspond to the labeled peaks on the ESI-MS spectra in Fig. 11. The proposed structure for each peak is given in the column on the left. Each structure differs from the one below it by the addition of either and Pentose or a HexNAc.

Exoglycosidase digestion and lectin binding to reveal that N-Acetylglucosamine (GlcNAc)is the terminal sugar of the disaccharide repeat.

In an attempt to identify the specific N-acetylhexosamine as well as the order and orientation of monosaccharides within the O-repeat, glycosidase digestion was performed. The results of the digestion are shown in Fig 13. β-N-Acetyl-hexosaminidase is an enzyme that is capable of cleaving terminal D-galactosamine (D-GalNAc) and D-glucosamine (D-GlcNAc) residues that are linked by β (1 – 2, 3, 4, 6) linkages to oligosaccharide chains. It has been previously used in our lab to characterize LOS produced by *N. gonorrhoeae* PID2 (127). The purified 4320 LOS and LPS mixture was incubated in the presence of this enzyme, and the resulting products were electrophoresed on an SDS-PAGE gel and silver stained alongside the undigested control molecules. The purified undigested *N. sicca* 4320 LOS and O-repeat is shown in lane 2. The prominent band with the fastest mobility is the LOS band, and the LPS containing the O-repeat bands can be seen to form a ladder on the gel. Drawing from the MALDI-MS data, each O-repeat band is larger than the one below it by one disaccharide repeat unit. Lane 1 shows the profile resulting from the purified molecules incubated in the presence of β-N-Acetyl-hexosaminidase. When lanes 1 and 2 are compared, a shift in mobility is evident. The O-repeat bands in lane 1 show an increase in mobility by one monosaccharide compared to the respective control band in lane 2. This enzymatic digest supports that HexNAc is the terminal sugar linked to the O-antigen by a β bond and that this monosaccharide is either D-GalNAc or D-GlcNAc.

To determine the specific identity of the HexNAc unit, lectins were employed that have affinity for specific monosaccharides. Biotinylated Succinylated Wheat Germ Agglutinin (WGA) (Vector Laboratories, Burlingame CA) binds solely to GlcNAc

residues. This lectin was used to probe 4320 LOS and LPS in Fig. 14. As a negative control, *N. sicca* 342 LOS was purified and electrophoresed on the SDS-PAGE gel alongside *N. sicca* 4320 LOS and LPS. *N. sicca* 342 fails to produce LPS, but produces a similar LOS to 4320. As shown in Western blot, the LPS of *N. sicca* 4320 bound succinylated WGA, while the LOS molecules of both 342 and 4320 did not bind the lectin. Although the lectin affinity was relatively weak, these results suggest that the O-repeat unit contains GlcNAc. GC-MS is used in the next section in an attempt to confirm the lectin affinity results.

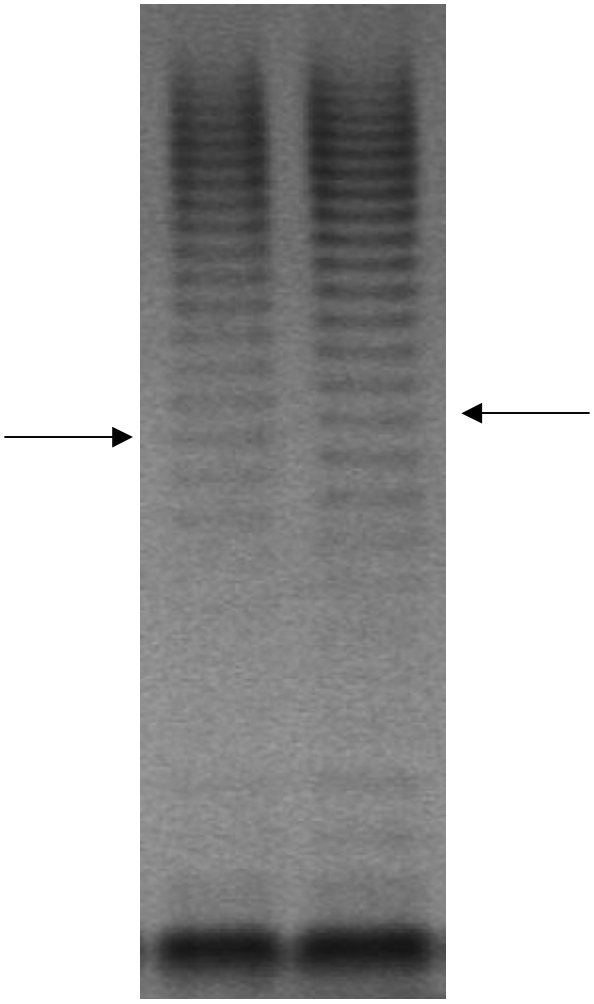


Figure 13. Exoglycosidase digestion of *N. sicca* 4320 LPS with N-Acetylhexosaminidase. Lane 1 shows the SDS-PAGE profile of *N. sicca* 4320 LOS and LPS after digestion with β-N-acetylhexosaminidase. Lane 2 contains the same undigested preparation. Preparations were run on a 16.5% Tris-Tricene polyacrylamide gel then silver stained. Arrows point toward bands demonstrating the shift in mobility that occurred upon addition of the glycosidase.

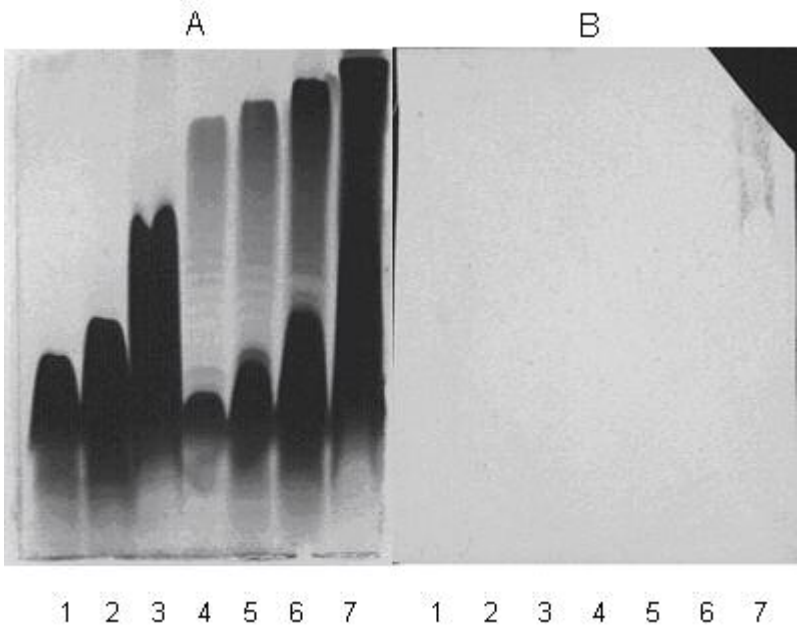


Figure 14. SDS-PAGE and lectin binding assay of *N. sicca* 342 LOS and *N. sicca* 4320 LOS and LPS. Panel A contains a silver stained SDS-PAGE gel containing 10-fold dilutions of *N. sicca* 342 LOS respectively in lanes 1 through 3. 10-fold dilutions of 4320 LOS and LPS are shown respectively in lanes 4 through 7. Panel B shows a Western blot of an identical gel with biotinylated succinlated WGA.

GC-MS analysis demonstrating that rhamnose and GlcNAc compose the disaccharide repeat.

To confirm that the *N. sicca* 4320 disaccharide repeat contains GlcNAc and to determine the identity of the other monosaccharide in the unit, GC-MS was performed by Dr. Vern Reinhold's laboratory at the University of New Hampshire. Figure 15 shows GC-MS of the L-rhamnose and D-GlcNAc controls. L-Rhamnose had a 11.10 minute retention time and generated a characteristic fragmentation pattern. The same procedure was performed for L-fucose (data not shown). The D-GlcNAc control, had a retention time of 24.51 minutes and was fragmented to show a separate mass spectral pattern. 4320 LOS and LPS was hydrolysed to monosaccharide components and the results are contained in Fig 16. This sample was composed of monosaccharides with 11.09 and 24.51 minute retention times. Upon fragmentation of these major peaks during mass spectral analysis, patterns were created that matched the control L-rhamnose and D-GlcNAc controls. The data from the MALDI-MS, ESI-MS, glycosidase digestion, lectin binding, as well as the GC-MS establish that the 4320 O-repeat is a disaccharide composed of D-GlcNAc and L-rhamnose.

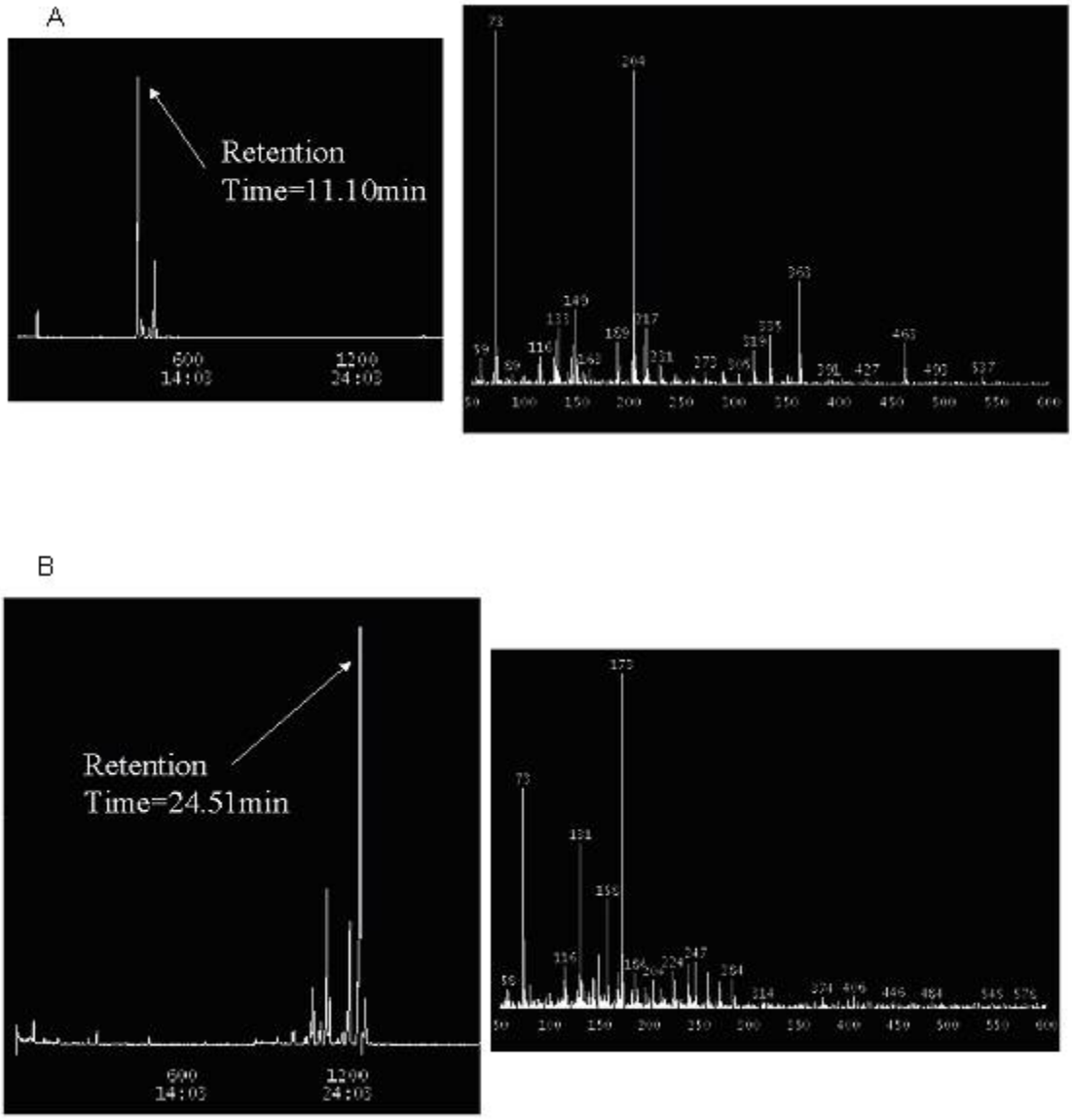


Figure 15. GC-MS analysis of L-Rhamnose and GlcNAc monosaccharide controls. Panels A and B respectively contain the profiles of the L-Rhamnose and D-GlcNAc monosaccharide controls. The retention times of these sugars along with the fragmentation pattern of the ion represented by the peak are shown. These figures were generated by Dr. Vern Reinhold's laboratory at the University of New Hampshire.

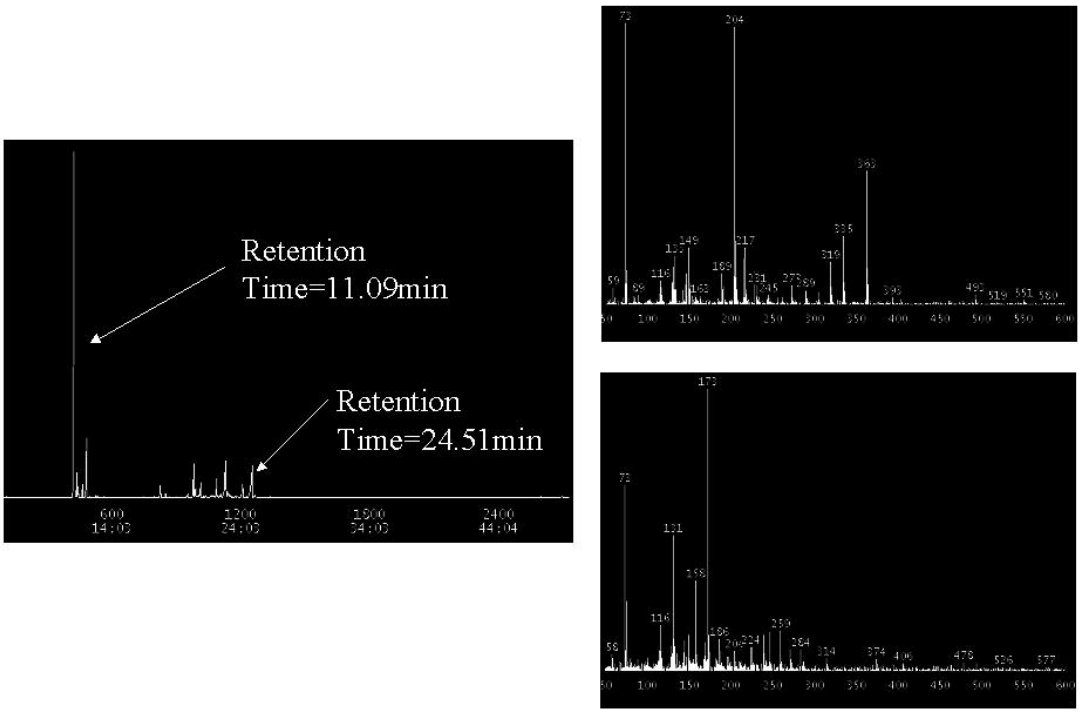


Figure 16. GC-MS analysis of *N. sicca* 4320 LPS. The panel on the left shows the retention time of the monosaccharide sugars contained in 4320 LOS and LPS. The retention times of the major peaks are shown. The right panels show the fragmentation pattern of the sugars represented by the peaks 11.09 and 24.51. These figures were generated by Dr. Vern Reinhold's laboratory at the University of New Hampshire.

Linkage analysis of disaccharide repeat.

The 4320 O-antigen is a disaccharide repeat composed of L-Rhamnose and D-GlcNAc. Therefore, two types of bonds exist in the polysaccharide. MS^n analysis was performed by Dr. Vern Reinhold's laboratory at the University of New Hampshire to identify the type and orientation of bond that connects D-GlcNAc to the reducing carbon of L-Rhamnose. ESI-MS was performed on non-pyrazole 4320 LOS and LPS with the spectra shown in Fig. 17. As in the pyrazole derivitized sample, fragments are seen with mass differences corresponding to the loss of methylated L-Rhamnose (174 *m/z*) and D-GlcNAc (245 *m/z*). For example, the ion corresponding to the 1745.9 *m/z* peak contains an additional GlcNAc compared to the 1572.1 *m/z* molecular ion and contains an additional disaccharide repeat compared to the 1326.9 *m/z* ion.

The ion represented by the 908 *m/z* peak was chosen as a starting point for MS^n analysis because it contains two methylated L-Rhamnose – D-GlcNAc repeats. During this analysis the 908 *m/z* ion was successively fragmented into smaller structures. Expected fragments and their masses are shown in Fig. 20. The MS^2 spectra showing the fragmentation of the 908 *m/z* ion is shown in Fig 18a. A 701 *m/z* major peak was seen, representing a loss of a methylated rhamnose from the parent ion. Fragmentation of the 701 *m/z* ion in shown in the MS^3 spectra yielded a major ion of 474 *m/z*, consistent with the loss of a methylated GlcNAc. Continuation of this pathway is shown in Fig. 18b. The MS^4 spectra shows that the 474 *m/z* ion generated the 282 *m/z* ion, leaving just methylated GlcNAc. Fragmentation of the 282 *m/z* ion resulted in the MS^5 spectra in which a major peak of 180 *m/z* was seen. As shown in Fig. 19, this cleavage pattern is

characteristic of a β (1-3) linkage to GlcNAc. This data shows that L-Rhamnose is connected to the 3 carbon of GlcNAc by a β bond.

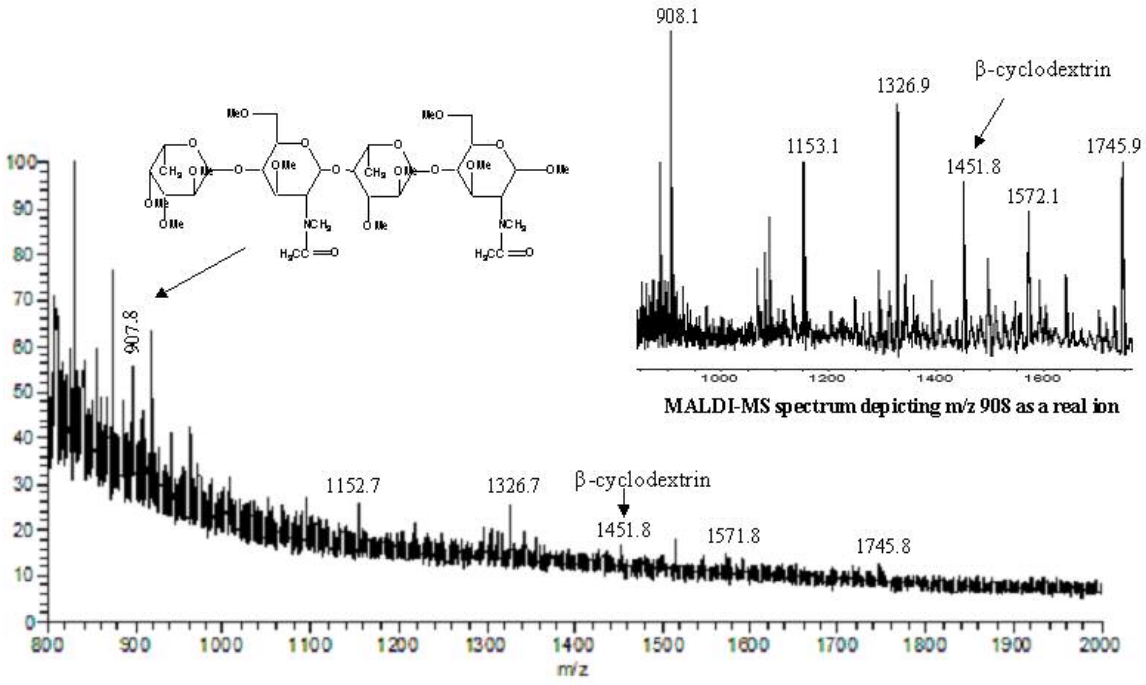


Figure 17. ESI-MS showing the generation of the starting molecule structure for MS^n analysis of *N. sicca* 4320 LPS. This spectra was generated by Dr. Vern Reinhold's laboratory at the University of New Hampshire.

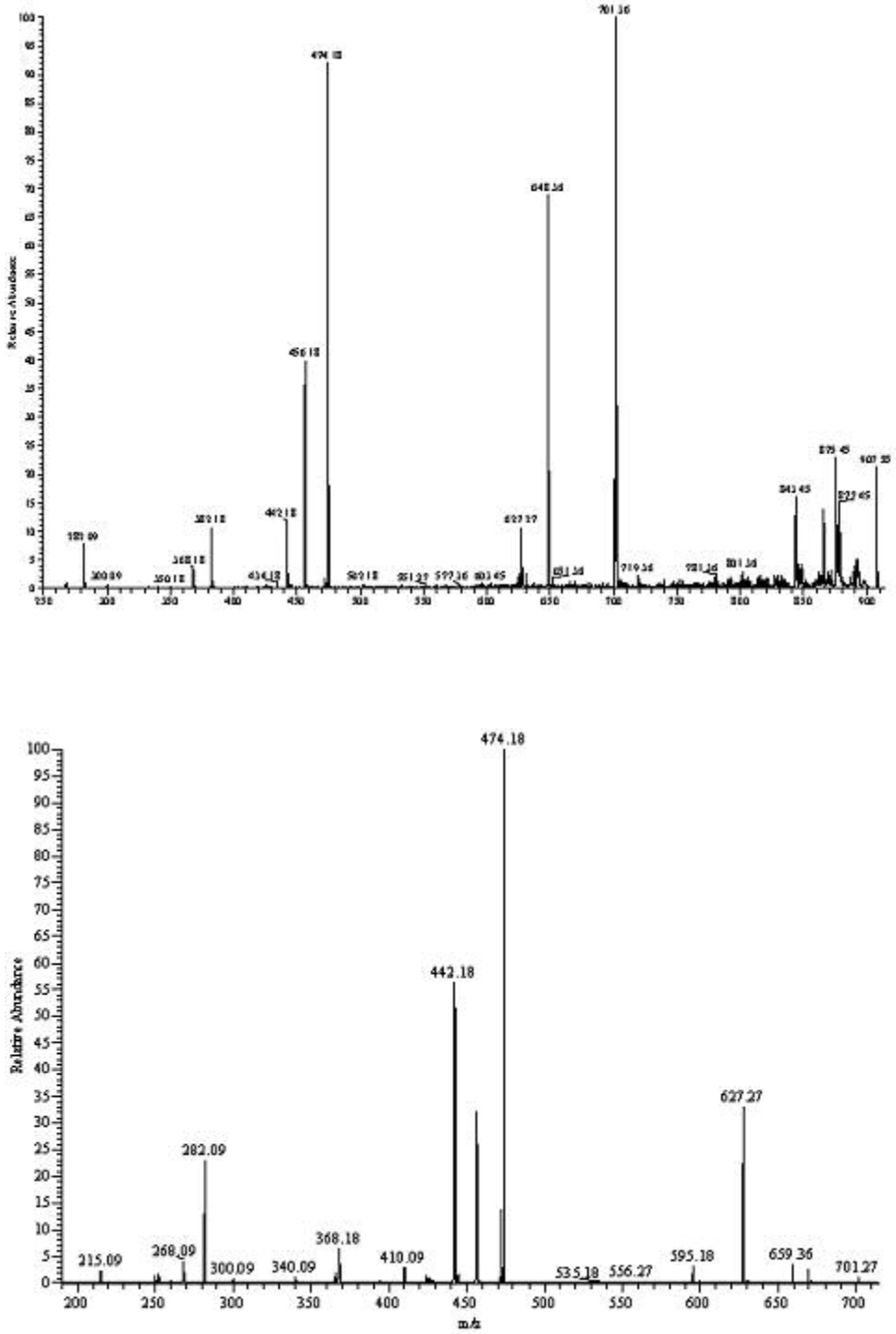


Figure 18a. MS^2 and MS^3 spectra from MS^n analysis of *N. sicca* 4320 LPS.

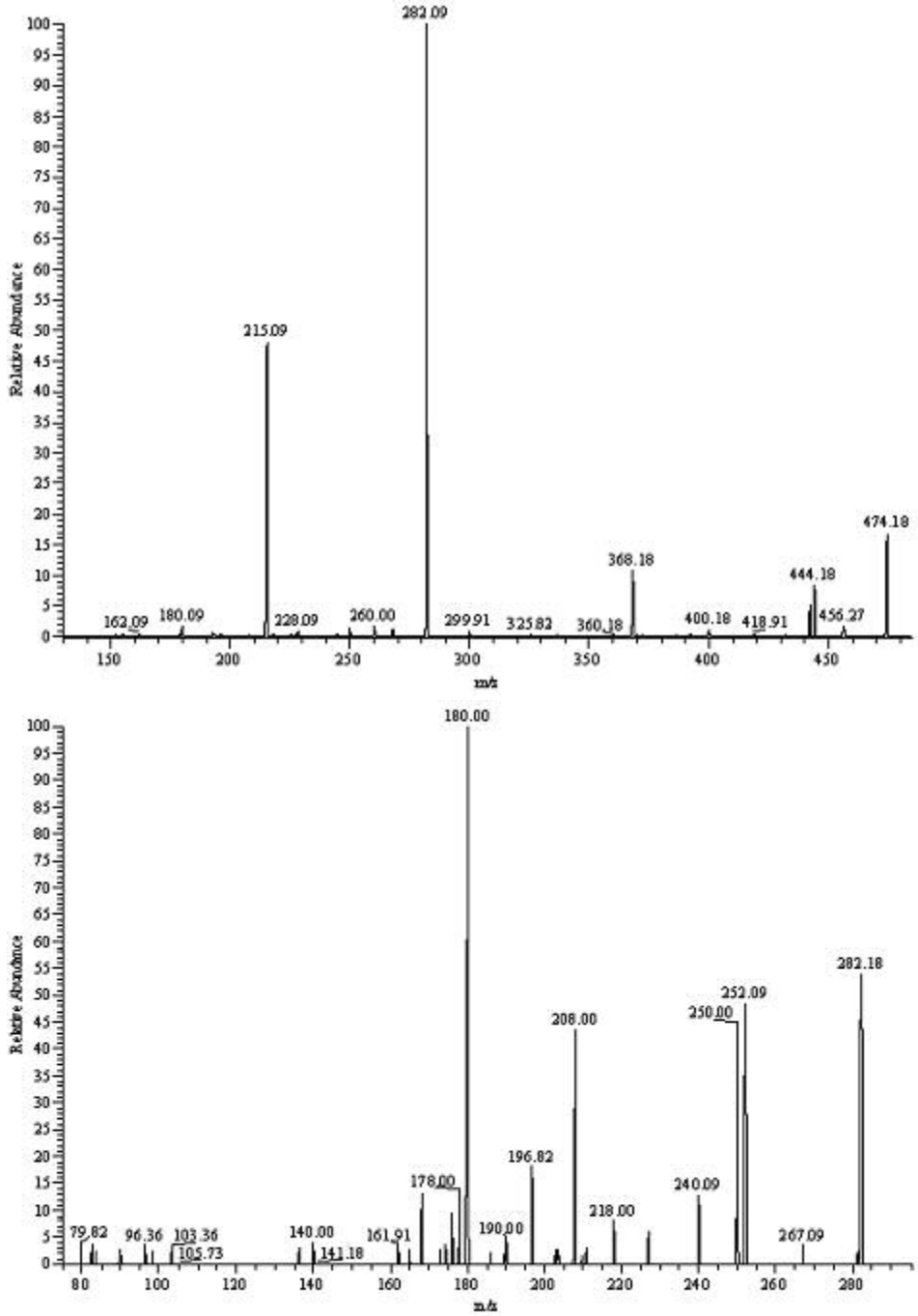


Figure 18b. MS^4 and MS^5 spectra from MS^n analysis of *N. sicca* 4320 LPS.

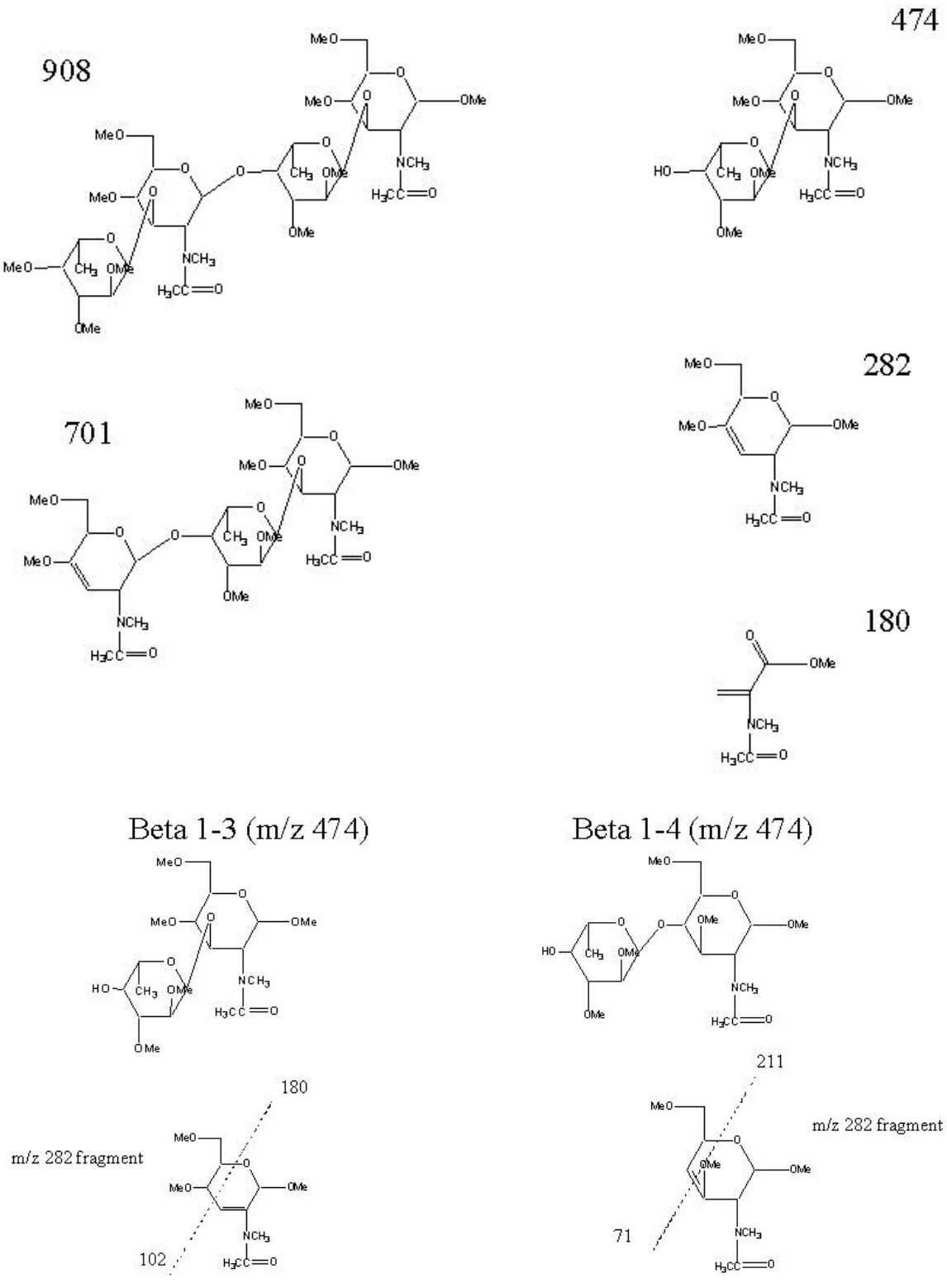


Figure 19. Proposed structures generated during MS^n analysis.

DISCUSSION

The purpose of this study was to identify the specific LPS structure that is expressed by *N. sicca* 4320. It was predicted that the LPS structure would be composed of polylactosamine, as this is the repeat comprising the *Neisseria* and *Haemophilus* LPS in the few instances that have reported the synthesis of this molecule. However, the combination of structural determination methods, MALDI-TOF, exoglycosidase digestion, lectin binding, GC-MS, and MS^n analysis, show that 4320 LPS is a different molecule from what was predicted.

The data demonstrates that *N. sicca* 4320 expresses two separate molecules. MALDI-MS of the LOS molecule is consistent with a backbone structure of $(Hex)_2(Hep)_3(PEA)_1(KDO)_1$(lipoidal moiety). This LOS structure is significant because it is typical of the molecule expressed by the *Haemophilus*, in that it contains three heptose residues in the inner core (45,96). *Neisseria* and *Haemophilus* are both naturally competent and known to inhabit the same mucosal niches. Therefore, it is not unreasonable that surface structures necessary for adherence to host cells would be similar in the two genus.

The data shows that the LPS is a separate molecule from LOS, as the O-antigen repeat exist at masses less than that of lipid A of LOS. This O-antigen is composed of a polysaccharide component composed of a dissacharide repeat of [– 3) β-D-GlcNAc (1-X) β-L-Rham (1-] with a GlcNAc reducing terminus. Analysis of the MALDI-TOF data shows that this O-antigen is not anchored to lipid A. Instead it is likely anchored to a membrane component that is either 504 or 155 mass units. This is a novel structure in

both the composisition of the O-repeat as well as the anchor component that has not been previously reported in the *Neisseria*.

A similar structure has been reported in the gram-negative pathogen, *Agrobacterium tumefaciens*. This species is known to cause crown gall disease in plants. The bacteria enters the wound site and adheres to the plant cell wall. This interaction is mediated by the O-repeat antigen of the LPS as well as the receptor proteins of the the cell wall (25). It has recently been discovered that *A. tumefaciens* F/1 produces two O-antigens, [-3)-α-L-Rhap-(1-3)-β-D-GlcpNAc-(1-] and the second one [-4)-α-L-Rhap-(1-3)-β-D-GlcpNAc-(1-]. However, these O-antigens are anchored to Lipid A (24).

The presence of the form of LPS determined in this study has yet to be described in the *Neisseria*. Further study of the synthesis of this molecule by *N. sicca* 4320 is warranted because it shows that the *Neisseria* possesses additional genetic mechanisms involved in LPS synthesis. For example, because the O-antigen does not extend from LOS, but rather from an unknown anchor additional proteins are likely required for this addition. Also, the presence of rhamnose in the O-repeat indicates that 4320 possesses rhamnose biosynthesis genes. The gonococcus is known to contain a cryptic rhamnose biosynthetic cluster (107). It is possible that 4320 represents a strain that is able to produce rhamnose from a fully functional rhamnose biosynthetic cluster. The genetics of this LPS production need to be analyzed to determine the capacity of the *Neisseria* to express LPS.

CHAPTER 5. DEVELOPMENT OF A TRANSFORMATION METHODFOR *N. SICCA* 4320

INTRODUCTION

Neisseria can exchange genetic information through a horizontal transfer process. This is one method by which polymorphism occurs in the species (35). This natural competence is dependent on type IV pilus machinery (1,138). Prepilin proteins are synthesized and transferred to the inner membrane where the protein complex PilDFGT assembles the PilE monomers into a pilus polymer. Addition of PilE to polymer results in elongation of the pilus through the PilQ complex of the outer membrane (40).

DNA containing the neisserial uptake sequence (NUS), 5'-GCCGTCTGAA, can bind to proteins ComP and ComE that are associated with PilE at the outer membrane (1,46,138). DNA is brought into the cell by force dependent PilT mediated retraction of the pilus (82,138). In the periplasm, DNA is escorted by proteins Tpc and ComL to the inner membrane (38,39). At the inner membrane, the DNA interacts with ComA to obtain access to the interior of the cell (33).

Neisseria have been shown to be transformed by either single stranded or double stranded DNA with equal effieciency (122). Upon the addition of double stranded DNA, both single and double stranded DNA were found in the periplasm (20). It is thought that in the periplasm RecBCD unwinds the double stranded molecule, leaving the single stranded form to bind to Tpc and ComL for transport to the inner membrane (87). Once the single stranded form reaches the interior of the cell, RecN and Rep helicase position the DNA for homologous recombination into the neisserial chromosome (75,117).

Neisseria that lack expression of the Type IV pilus are poorly transformable. Other techniques have been employed to transform non-piliated *Neisseria*. One study

attempted to transform non-piliated gonococcal cells by electroporation (42). Gonococci were successively washed with poly-ethylene glycol 5000 (PEG-5000) while at 4^{o}C. DNA was added to the cells prior to the delivery of an electric pulse. A method based on chemical competency was developed to transform non-piliated *N. meningitidis*. Cells were incubated with calcium chloride for fifteen minutes at 4^{o}C. DNA was added and the cells were heat shocked (15). Although these techniques resulted in the transformation of *Neisseria*, the transformation frequencies were far below the 10^{-2} to 10^{-3} that is typically seen when naturally competent piliated cells are transformed. Therefore, these techniques were not appropriate for use in generating a collection of random transposon mutated cells. In addition, these two methods demonstrate mixed results when applied to various strains. My initial results at 4320 tranformation by natural competency proved unsuccessful, so a transformation method was adapted for *N. sicca* 4320. The application of the Inoue chemical competency method resulted in efficient transformation of 4320 at a frequency required for the generation of random 4320 mutants. These random mutants could then be screened for cells expressing an altered LPS.

RESULTS

Construction of DNA to create random mutations in *N. sicca* 4320.

To disrupt genes in the *N. sicca* 4320 chromosome, random transposon mutagenesis was performed. Because transformation of the 4320 chromosome requires excessive amounts of DNA when compared to other *Neisseria*, mutant DNA needed to be in a form that was easily replicated. Therefore, a strategy was undertaken to clone 4320 chromosomal fragments into the vector pUC19, mutate these clones, and then replicate the mutant clones prior to transformation of *N. sicca* 4320.

Clones containing random 4320 chromosomal fragments were constructed. *N. sicca* 4320 chromosomal DNA was isolated and digested with EcoRI. Fragments ranging is size from 600 base pairs to greater than 9 kb were visible upon agarose gel electrophoresis. The vector pUC19 was also digested with EcoRI and revealed a single band of the expected size, 2.7 kb. The 4320 chromosomal fragments were ligated to pUC19 and the resulting constructs were transformed into *E. coli* and screened on LB media containing ampicillin and X-gal. White ampicillin resistant colonies were assumed to contain the 4320 clones. Plasmids were extracted from individual colonies and digested with EcoRI, as shown in Fig. 20. Each clone contained the predicted 2.7 kb pUC19 digestion product along with an insert, indicating that the white *E. coli* ampicillin resistant colonies contained random clones. To statistically ensure that all of the chromosomal DNA was represented upon combination of the plasmids, 5416 colonies were obtained.

Once clones were constructed that contained a random sampling of the complete 4320 chromosome, the plasmids were pooled. As shown in Fig. 21 when the pooled

random clones were digested with EcoRI the characteristic 2.7 kb pUC19 digestion product could be seen compared to the digested pUC19 control. However, digestion of the pooled product also revealed the *N. sicca* 4320 chromosomal fragments of size varying from 600 bp to greater than 9 kb.

The recombinant plasmids were used as a target for transpososon mutagenesis. The pooled plasmids were reacted *in-vitro* with Epicentre's EZ::Tn kit. The reaction product was used to transform *E. coli* DH5α and kanamycin-resistant transformants were isolated. Over 10,000 resistant colonies were collected from the kanamycin media and plasmids were extracted. As depicted in Fig. 21 when the transposon mutant plasmid pool was digested with EcoRI, the plasmids still contained the 4320 chromosomal inserts between 600 bp and greater than 9 kb. Transposon insertion was also seen as expected in the mutant clones. The digestion product contains two prominent bands, one at 2.7 kb and one at 3.9 kb which are pUC19 and pUC19::Tn5 respectively. This shift in size is consistent with the insertion of a 1.2 kb transposon. The resulting mutant plasmid pool now contained random transposon insertions into *N. sicca* 4320 fragments and was easy to replicate to provide the excessive DNA needed for transformation.

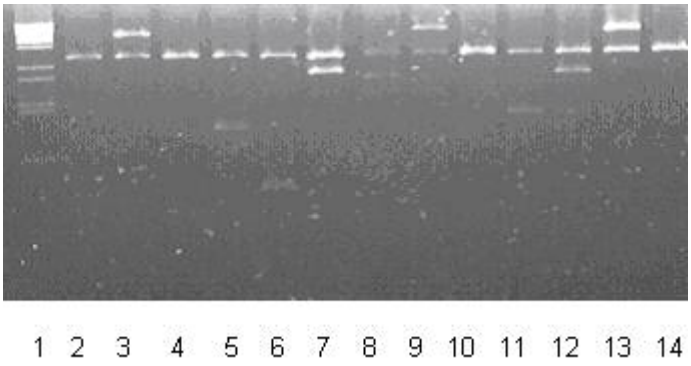


Figure 20. EcoRI digestions of random *N. sicca* 4320 shotgun clones. Lane 1 shows the λ BstEII standard. Lanes 2 and 14 contain pUC19 digest with EcoRI. The other lanes show EcoRI digested plasmids that were extracted from individual white ampicillin resistant *E. coli* transformants.

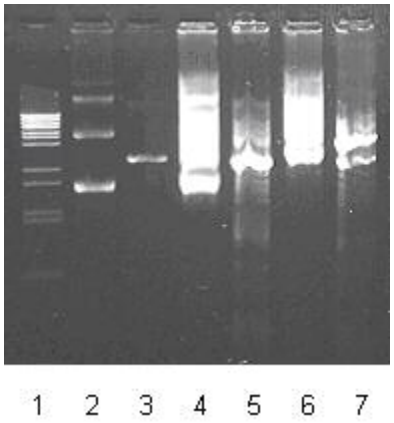


Figure 21. Transposon mutagenesis of random *N. sicca* 4320 clones. Lane 1 shows the λ BstEII digested standard. Lanes 2 and 3 respectively show pUC19 and its EcoRI digestion product. Lanes 4 and 5 contain the plasmids extracted from the 5416 random clones and the EcoRI digestion product of these clones. Lanes 6 and 7 show the shift in mobilities that occurred when the transposons were inserted into the plasmid pool and this pool was digested with EcoRI.

Creation of competent *N. sicca* 4320 through treatment with $CaCl_2$ followed by heat shock.

Chemically competent *N. sicca* 4320 cells were created in an analogous fashion to the Inoue *E. coli* method (60). *N. sicca* 4320 was grown overnight on GCK media and a single colony was selected and used to inoculate a 25 ml GCP broth containing $NaHCO_3$ and Kellogg's supplement. The culture was incubated at 37°C with shaking at 250 rpm for 6-8 hours. The broth culture was then used to inoculate 250 ml of GCP also containing $NaHCO_3$ and Kellogg's supplement. The flask was incubated with shaking overnight at 25°C.

The following morning, the OD_{600} of the flask was taken and repeated every hour until an OD_{600} of 0.55 was reached. If needed, the flask was moved to a temperature of 37°C in order to obtain a culture with the required optical density. Next, the flask containing the *N. sicca* 4320 cells was placed in an ice bath for 10 minutes. The cells were centrifuged at 3900 rpm for 10 min at 4°C. The supernatant was decanted and the pelleted bacteria was placed on ice. Excess GCP was removed from the pellet prior to the addition of 80 ml of ice cold Inoue transformation buffer. Cells were resuspended in the buffer by swirling the bottle while remaining on ice. Cells were centrifuged at 3900 rpm for 10 minutes at 4°C. The supernatant was removed and the pelleted bacteria placed on ice. Transformation buffer (20 mL, 4°C) was added and the 4320 cells were gently resuspended. To the suspension, DMSO (1.5 mL) was added and the resulting solution was stored on ice for 10 min. Aliquots were made in microcentrifuge tubes and immediately snap frozen in a dry ice-ethanol bath prior to storage at –80°C.

N. sicca 4320 cells were transformed by the heat shock method. Competent cells were removed from –80°C and thawed on ice for 10 minutes. Transposon-mutated

plasmid DNA was added to the cells (15 μl of 100 ng/ml of DNA) and kept on ice for 10 minutes. The mixture was shocked at 37^{o}C for two minutes and 900 μl of room temperature GCP with growth supplements was added. Cells were grown for one hour at 37^{o}C prior to plating on GCK media containing kanamycin.

After overnight incubation, the transformants were counted and yielded a transformation frequency of 2.34 X 10^{-6}. Transformants were selected at random for colony PCR analysis to confirm the presence of the transposon. The PCR products from this analysis are shown in Fig. 22. Primers specific for the amplification of the kanamycin resistance marker were used and the transformants yielded an amplification product of the desired size, 1.2 kb. The same reaction using 4320 colonies as a template failed to produce any amplicon.

To test the ability of *N. sicca* 4320 to be transformed, chromosomal DNA from one of the *N. sicca* 4320 kanamycin resistant transformants was isolated. This DNA was used to transform the competent 4320 cells according to the previously described heat shock protocol. After incubation, colonies were again seen on the selective media at a frequency of 6.2 X 10^{-6}.

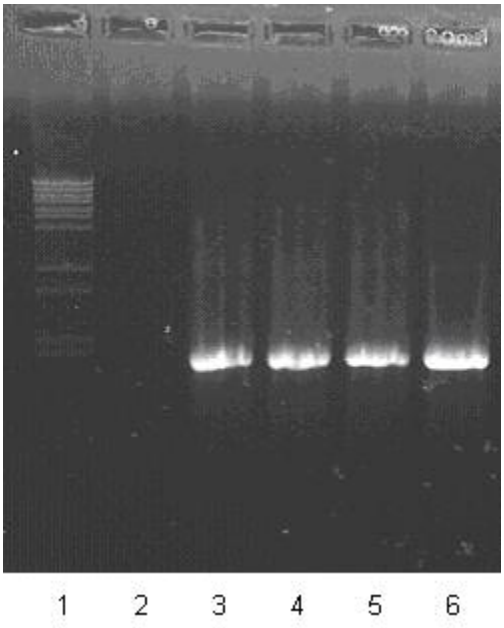


Figure 22. Colony PCR of *N. sicca* 4320 transformants. Lane 1 contains the λ BstEII digested standard. Lanes 2 through 6 show the resulting PCR products that were amplified from primers Kan-F and Kan-R which were designed to amplify the kanamycin resistance marker. Templates included *N. sicca* 4320 (lane 2), kanamycin resistant *N. sicca* 4320 transformants (lanes 3-5), and the transposon from the Epicentre kit (lane 6).

Generation of random mutations in *N. sicca* 4320.

To create random mutations in the *N. sicca* 4320 chromosome, the mutant plasmid pool was used to transform 4320, according to the previously described protocol, and kanamycin resistant colonies were selected. To ensure that transposon insertions occurred randomly throughout the chromosome, a Southern blot was performed.

Sequence analysis of the transposon construct was undertaken to determine a restriction digestion strategy that would show insertions at differing sites within the chromosome. The enzyme BstEII cleaves the transposon at two sites. When inserted into the chromosome, digestion with BstEII would cleave the transposon into three fragments. One fragment would be 300 bp, the distance between BstEII sites in the transposon. The other two fragments would be composed of the sequence between the BstEII site in the transposon and the closest BstEII site in the DNA flanking the transposon. Therefore, the size of the later two fragments is determined by the site of integration in the chromosome.

Ten random *N. sicca* 4320 kanamycin resistant transformants were selected for the Southern blot analysis. Chromosomes from the mutants were isolated along with the chromosomes from two strains, *N. sicca* 4320 and *N. gonorrhoeae* F62ΔLgtAlpt3::Tn5. The former strain was chosen to act as a negative control, as it did not contain a transposon insertion. The later strain was known to have a transposon insertion in *lpt3*. As shown in Fig. 23., transposon insertion occurred in a random pattern in the 4320 chomosome. As expected the kanamycin resistance probe hybridized to 300 bp in all 10 chromosomes. Also, hybridization can be seen to two variable bands that demonstrate the transposon is located at a different site in each of the selected chromosomes. Additional evidence supporting the integration of the transposon into the 4320 chromosome is

provided by the failure to isolate plasmid DNA from the 4320 kanamycin resistant transformants, as well as the failure of the chromosomal DNA from the 4320 transformant to transform *E. coli*.

Genetic transformation of *N. sicca* 4320 is likely mediated by a double recombination event. If a single crossover recombination event occurred, pUC19 sequence would have recombined into the chromosome along with the transposon mutagenized sequence. However, this event did not take place because a Southern blot probed with DIG labeled pUC19 was performed on an identical gel and failed to show the expected hybridization signals.

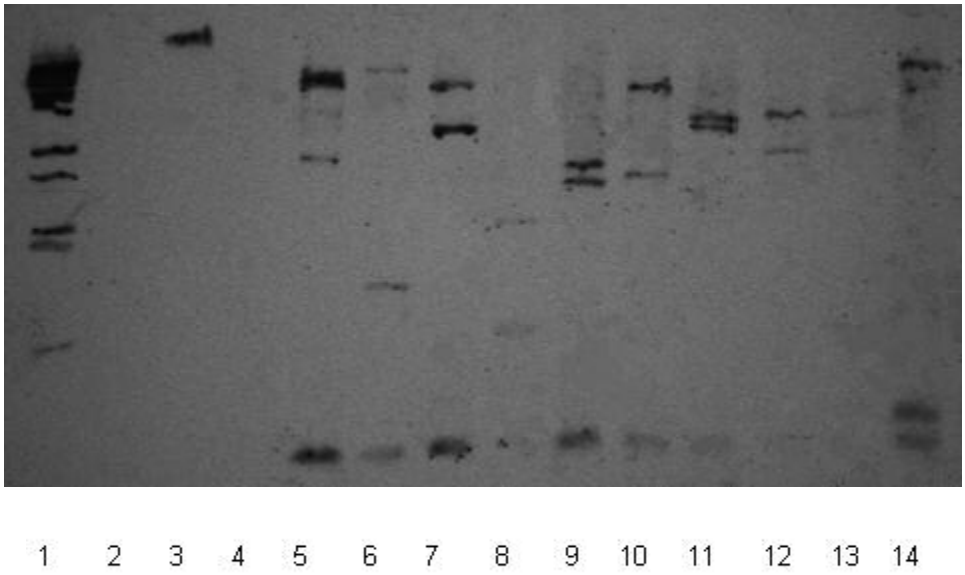


Figure 23. Southern hybridization of random *N. sicca* 4320 transformants. Lane 1 contains DIG labeled λ BstEII standard. Lanes 2 and 3 respectively contain the undigested chromosomes from the negative and positive controls, *N. sicca* 4320 and *N. gonorrhoeae* F62ΔLgtAlpt3::Tn5. Lane 4 contains the BstEII digested 4320 chromosome. Lanes 5 through 14 show the hybridization pattern generated when 10 random 4320::Tn5 chromosomes were digested with BstEII and probed with a DIG labeled kanamycin resistance marker.

Screening for *N. sicca* 4320 LPS mutants.

In order to identify 4320 transformants exhibiting an altered LPS it was necessary to create a reagent that would differentiate between *N. sicca* 4320 expressing wild type LPS and mutant LPS. One way of developing this type of reagent was to generate Ab that was specific for the LPS molecule. Therefore, failure of a 4320 transposon mutant to bind the Ab would indicate that the specific transformant fails to express the wild type LPS. Like 4320, *N. sicca* 4319 expresses LPS. Membrane blebs from this strain were previously used to generate polyclonal Abs. In this study, these antibodies were absorbed against *N. sicca* 342, a strain that does not synthesize LPS but produces an LOS with similar SDS-PAGE mobility as 4320. The absorption process is depicted in Fig. 24. Prior to absorption against 342, the mixture of antibodies bound to both *N. sicca* 342 and *N. sicca* 4320. After these Abs had been incubated with 342 cells, an identical colony blot was performed. 342 cells no longer reacted with the mixture, while 4320 cells still reacted. It was then hypothesized that the antibodies that were not absorbed by 342 cells would be reactive with the 4320 LPS.

To demonstrate that the Ab that remained after absorption still reacted with the *N. sicca* 4320 LPS, a Western blot was performed. LOS that was extracted from *N. sicca* 342 was electrophoresed on an SDS-PAGE gel alongside the 4320 LPS. The gel was silver stained to visualize the molecules as shown in Fig. 25. Different dilutions of each preparation were added to the gel to determine the sensitivity of the antibody. However, the LOS bands of 342 and 4320 can readily be seen on the gel as well as the 4320 LPS bands. An identical gel was used for a Western blot with the absorbed antibody. The results of the blot are also exhibited in Fig. 25. The LOS of both the 342 and 4320 strains

failed to react with the remaining antibody while the bands composed of the Rhamnose-GlcNAc disaccharide repeat showed reactivity. Therefore, absorption of the antibodies generated against *N. sicca* 4319 with *N. sicca* 342 cells resulting in the collection of antibodies that lost reactivity with 4320 LOS, but remained active against the LPS.

It was believed that the LPS reactive antibody that was generated by the absorption process would be useful to screen 4320 transposon mutants for a change in LPS phenotype. Upon transformation with the pool randomly mutagenized 4320 chromosomal fragments, kanamycin colonies were selected and patch plated in duplicate onto GCK media supplemented with kanamycin. In total, 5000 colonies were transferred to a nitrocellulose membrane that was used for a colony blot with the absorbed antibody. Upon development of the blot, all colonies retained the ability to react with the antibody. Colonies which were slightly less reactive with the antibody when compared to the wild type *N. sicca* 4320 control were selected for further analysis. Upon electrophoresis on an SDS-PAGE gel, the LPS of the less reactive class did not display and altered LPS.

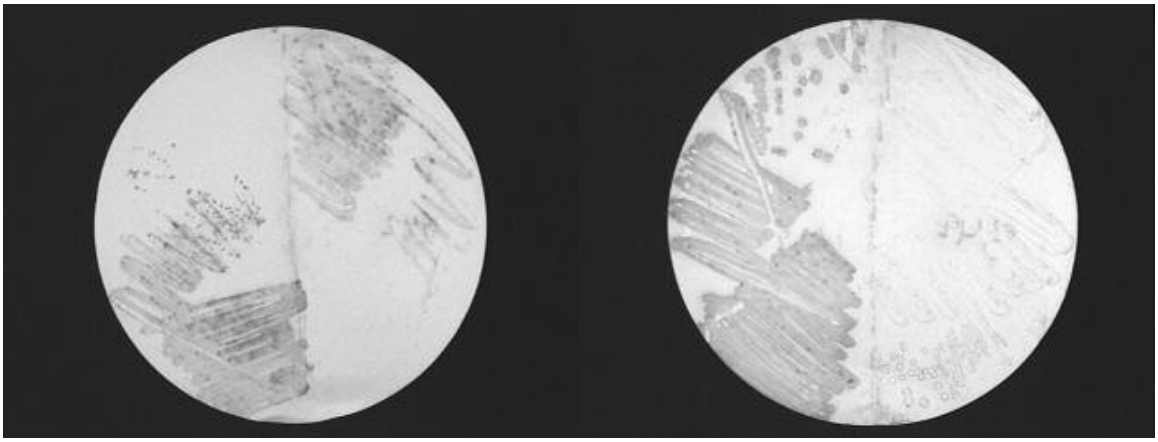


Figure 24. Colony blot of *N. sicca* 342 absorbed antibody. The left blot shows antibody that was produced against *N. sicca* 4319 whole cells in reaction with *N. sicca* 4320 (left side) and *N. sicca* 342 (right side). The blot on the left shows a blot of similar orientation that was reacted with antibody after the absorption process.

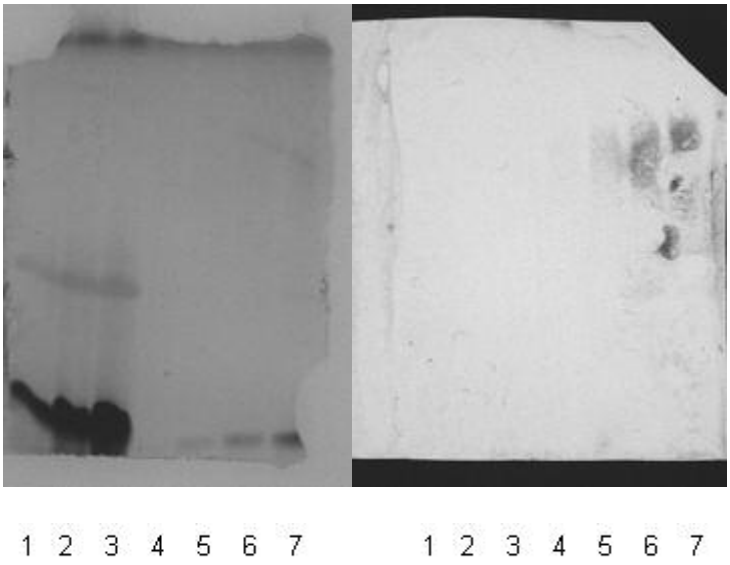


Figure 25. SDS-PAGE and Western blot of *N. sicca* 4320 LOS and LPS with *N. sicca* 342 absorbed antibody. Panel A (right) shows a silver stained SDS-PAGE gel containing 10-fold dilutions of *N. sicca* 342 LOS (lanes 1-3) and *N. sicca* 4320 LOS and LPS (lanes 4-7). Panel B (left) shows a Western blot of an identical gel with antibody the obtained after absorption of *N. sicca* 4319 antibody with *N. sicca* 342 cells.

Since the Ab absorption screen failed to identify a *N. sicca* 4320 LPS mutant, other strategies were adopted. Another method was to utilize the LPS structural information that was generated in chapter 2 of this work. For example it was known that the disaccharide repeat contained GlcNAc. As described in Fig. 14, succinylated WGA lectin, which is specific for GlcNAc, was able to bind to the O-repeat bands of *N. sicca* 4320 LPS, but not to the LOS molecule. Therefore, it was thought that it would be able to discriminate between 4320 wild type and mutant LPS through a lectin binding assay. As in the colony blot with the absorbed Ab, *N. sicca* 4320 random transposon mutants were patch plated onto duplicate GCK plates that were supplemented with kanamycin. After transfer to a nitrocellulose membrane, the mutants were reacted with the lectin. Once again, after screening several thousand transformants, no LPS mutant was identified.

Another approach taken was to try to obtain a LPS mutant by randomly selecting over 1000 *N. sicca* 4320 transposon mutants. Upon random selection of each single transformant, a quick LPS preparation was made and was electrophoresed on an SDS-PAGE gel. No LPS molecules were altered from the wild type phenotype.

One final approach to identify a LPS mutant utilized the fact that rhamnose was contained in the disaccharide repeat. Although both the gonococci and the meningitidis do not include rhamnose in their LOS, both species are known to contain a cryptic cluster of rhamnose biosynthesis genes, *rfbBAD* (107). These genes are highly homologous to the rhamnose biosynthetic cluster in *E. coli*, an organism that incorporates rhamnose into its LPS. Therefore, it was reasonable that the rhamnose biosynthetic cluster could also be present in *N. sicca* 4320 and that inactivation of these genes could result in the alteration of the LPS phenotype.

To test for the presence of the *rfb* gene cluster in *N. sicca* 4320, PCR analysis was first performed. Primers based on the *N. gonorrhoeae* FA1090 sequence were designed to amplify each individual gene within the cluster. As shown in Fig. 26, when *N. gonorrhoeae* F62 chromosomal DNA was used as a template in conjunction with primers rfbB-F and rfbB-R a 1040 bp amplicon resulted that was consistent with *rfbB*. Reaction using primers rfbA-F and rfbA-R yielded a 866 bp product that was consistent with *rfbA*. When the F62 template was reacted with primers rfbD-F and rfbD-R a 986 bp product consistent with *rfbD* was seen. Also shown is the amplification product that resulted when the reaction was performed using primer rfbB-F and rfbD-R. The generated band was consistent with the entire *rfbBAD* gene cluster. To confirm that these PCR products were actually rhamnose biosynthetic genes, the amplicons were sequenced. Results showed that genes *rfbBAD* were amplified, respectively encoding the proteins dTDP-D-glucose 4,6-dehydratase, glucose-1-phosphate thymidylytransferase, and dTDP-6-deoxy-L-mannose-dehydrogenase. The same series of reactions were performed using the *N. sicca* 4320 chromosome as a template. However, upon electrophoresis amplicons of the characteristic sizes were not observed.

Because PCR analysis failed to detect the presence of the rhamnose biosynthetic gene cluster in *N. sicca* 4320, Southern hybridization studies were undertaken. PCR products from the previous experiment were known to be the *rfb* genes. These amplicons were labeled with DIG to create probes necessary for analysis. Three initial blots were performed, each one to detect the presence of each *rfb* gene. When compared to the *N. gonorrhoeae* F62 control, 4320 showed hybridization to probes *rfbB* and *rfbA*. However, no signal was seen with probe *rfbD*. The results of a more informative hybridization

experiment are shown in Fig. 27. *N. gonorrhoeae* F62 chromosomal DNA was separately digested with AgeI and EcoRI. Both enzymes were selected because they fail to cleave within the *rfb* cluster. Therefore, when probe *rfbBA* hybridized to the digested F62 only one signal was seen, indicating the linkage of the genes. The 4320 chromosome was separately digested with the same two enzymes and probed with *rfbBA*. Again, only one hybridization signal was seen for each digestion. This data indicates that *N. sicca* 4320 contains both *rfbB* and *rfbA* and that these genes are linked.

Since the hybridization results showed that there is similarity between the *N. gonorrhoeae* and *N. sicca* 4320 *rfbB* and *rfbA* genes, it was thought mutated gonococcal *rfbB* and *rfbA* could be used to inactivate these genes in the 4320 chromosome. Figure 28 depicts the process of constructing a mutation in the rhamnose biosynthesis genes. Primers rfbB-F-2-up and rfbA-R were used to amplify *rfbBA* while simultaneously adding a neisserial uptake sequence to the amplicon. The product was digested with EcoRI and cloned into pUC19 at its respective site. Once a clone, pUC19::*rfbBA* was obtained that yielded a 2.7 kb and a 1.8 kb band upon digestion with EcoRI, mutagenesis was performed. Random transposon mutagenesis was performed on the clone according to the Epicentre EZ::Tn protocol. After transformation of the reaction into *E. coli*, kanamycin resistant colonies were selected and mutant clones were extracted. EcoRI digestion of the mutated clones yielded a pattern that is characteristic of transposon insertion into both the pUC19 part of the clone and of the *rfbBA* portion. The mutated *rfbBA* construct was used to transform *N. sicca* 4320 according to the protocol previously described in this chapter. However, no transformants were obtained.

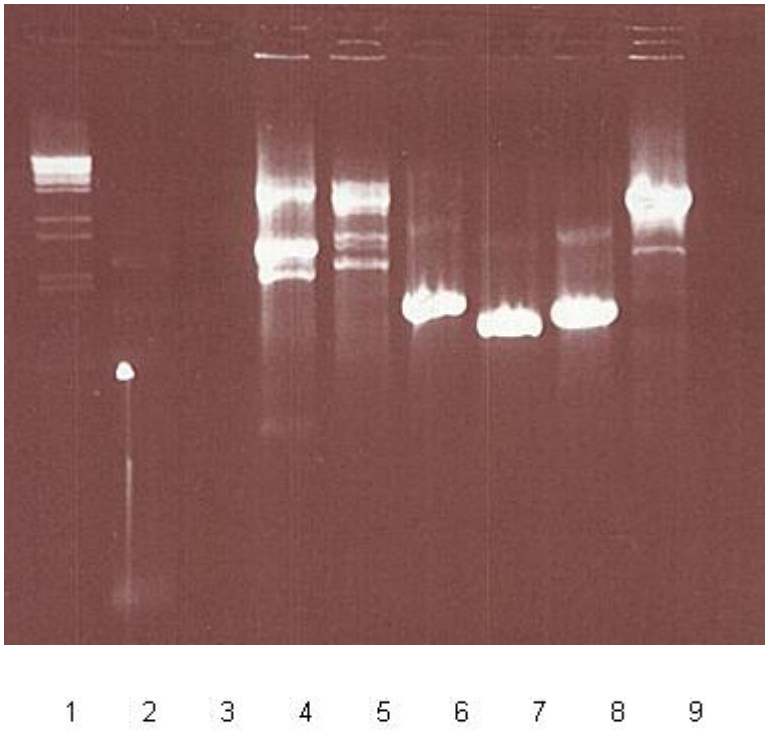


Figure 26. PCR analysis of rhamnose biosynthesis genes in *N. sicca* 4320. Lane 1 contains the λ BstEII standard. Lanes 2 through 5 show the PCR products yielded using the *N. sicca* 4320 chromosome as a template in reactions designed to respectively amplify *rfbB*, *rfbA*, *rfbD*, *rfbBAD*. Lanes 6 through 9 contain the same reactions using *N. gonorrhoeae* F62 chromosomal DNA as a template.

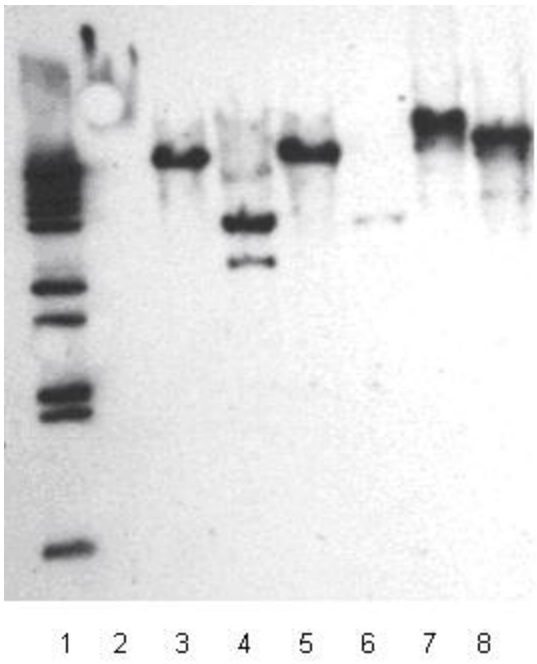


Figure 27. Southern hybridization of rhamnose biosynthesis genes. Lane 1 contains DIG labeled λ BstEII standard. Lane 2 contains undigested *N. sicca* 4320 chromosomal DNA. Lanes 3 and 4 contain *N. sicca* 4320 chromosomal DNA digested respectively with EcoRI and AgeI. Lanes 5 and 6 are duplicate lanes of 3 and 4. Lanes 7 and 8 show *N. gonorrhoeae* F62 DNA respectively digested with EcoRI and AgeI. Chromosomal DNA was hybridized with DIG labeled *rfbBA*.

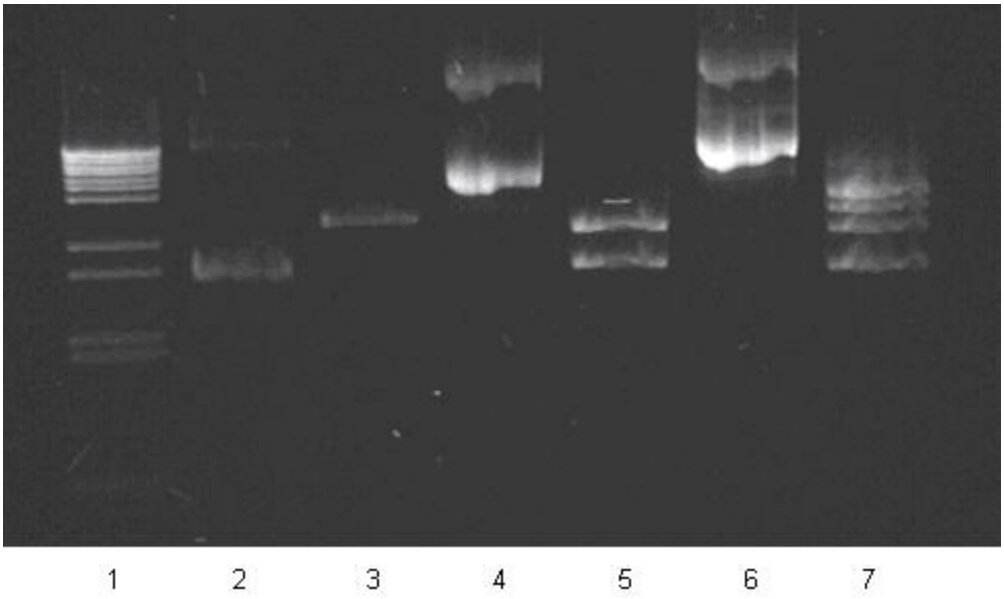


Figure 28. Construction of rhamnose biosynthesis mutant DNA. Lane 1 contains the λ BstEII standard. Undigested pUC19 is shown in lane 2 and is digested with EcoRI in lane 3. Lanes 4 and 5 contain pUC19::*rfbBA* and the clone digested with EcoRI. Lanes 6 and 7 show transposon mutated pUC19::*rfbBA* and the same sample digested with EcoRI.

DISCUSSION

In the previous chapter it was determined that *N. sicca* 4320 was capable of producing a LPS molecule consisting of a repeating rhamnose and N-acetylglucosamine dissacharide. To identify *N. sicca* 4320 LPS biosynthetic genes, random *in vitro* transposon mutagenesis of the chromosome was attempted using Epicentre's EZ::Tn kit. However, the 4320 cells were unable to be transformed with the randomly mutated chromosomal DNA by the natural competency method. It is likely that *N. sicca* 4320 is defective in the synthesis of the type IV pilus apparatus. The absence of any component involved in the binding of DNA or retraction of the pilus such as PilDFGT, PilE, PilC, PilQ, Tpc, ComL, or ComA can dramatically reduce the transformation frequency in the *Neisseria*.

In order to transform 4320, a procedure was developed that was adapted from the Inoue *E. coli* transformation method (60). The addition of calcium chloride to the chilled cells creates a pore in the membrane through which DNA can enter the cell. Increasing the temperature of the cells increases the fluidity of the membrane, resulting in the closing of the chemically created pores.

Evidence of transformation of the chemically competent *N. sicca* 4320 cells was seen in the results of several experiments. The addition of the random transposon mutagenized 4320 shotgun clones to the competent cells resulted in transformants that were resistant to kanamycin. Colony PCRs of these kanamycin resistant colonies showed the presence of Tn5 in each individual transformant, indicating that the mutant DNA pool had entered 4320. To discover if the mutant DNA had recombined into the chromosome or was present as a replicative plasmid in the cell, two preparations were done with the

transformants. First, plasmid extractions from the transformants were attempted but failed to yield any product. Since plasmid DNA could not be obtained from the cells, chromosomal DNA was collected. The isolated transformant chromosomes were used to transform the chemically competent kanamycin sensitive 4320 cells to kanamycin resistance. In addition, the isolated chromosomes failed to transform *E.coli*. This data shows that transformation of 4320 with the mutant plasmid pool resulted in the integration of the mutant DNA into the chromosome. To determine the mechanism by which the mutant DNA recombined into the chromosome, Southern blot analyses were performed. Chromosomes from kanamycin resistant transformants were digested with BstNI and probed with the kanamycin resistance gene contained within the Tn5 construct. These same chromosomes were also tested for their ability to hybridize with pUC19, the vector used to create the mutant plasmid pool. Because no pUC19 hybridization signal was observed while a random hybridization pattern was seen when chromosomes were probed with the kanamycin resistance marker, the mutant DNA must enter into the chromosome through a recombination mediated by a double crossover event.

After the development of a technique to transform *N. sicca* 4320, the chromosome of the organism could be manipulated in order to identify genes involved in LPS biosynthesis. It was thought that transposon insertion into a LPS biosynthetic gene would result in 4320 expressing an altered LPS phenotype. To screen for mutant LPS expressing mutants, antibody against LPS producing 4319 cells was incubated with 342 cells which do not express this molecule. Unabsorbed antibody was shown in Western analysis to be reactive against the 4320 disscharide repeat. Upon reaction with the absorbed antibody,

all transformants displayed specificity. Therefore, transformants displaying mutant LPS were not identified by this screen. One explanation for these results is that several differences might exist between *N. sicca* 342 and 4319. Proteins including Opa, Por, and Pilin exhibit antigenic variation across the *Neisseria* genus. If 342 displayed a different Opa protein then 4319, absorption of 4319 antibody with 342 will result in a mixture of antibodies with specificity for not only LPS, but Opa as well.

Another approach to engineer an LPS biosynthetic mutant was to directly mutate genes that showed homology to rhamnose biosynthesis genes. Southern hybridization analysis in this study showed hybridization the gonococcal cryptic rhamnose biosynthesis gene cluster to the 4320 chromosome. The gonococcus contains the genes *rfbBAD*, respectively encoding dTDP-D-glucose-4,6-dehydratase, glucose-1-phosphate-thymidyltransferase, and dTDP-6-deoxy-L-mannose dehydrogenase. It was thought that the absence of *rfbC*, encoding dTDP-6-D-glucose-3,5-epimerase, from this cluster results in the inability of the gonococcus to synthesize rhamnose. Two gonococcal genes that hybridized with the 4320 chromosome, *rfbB* and *rfbA* were cloned into pUC19 and mutated at random sites by transposon insertion. When this DNA was reacted with the competent 4320 cells, no transformants were observed. One explanation for the absence of transformants is that nucleotide diversity between the clusters in 4320 and *N. gonorrhoeae* F62 may exist, preventing recombination of gonococcal DNA into the 4320 chromosome. However, this explanation is doubtful as the bioinformatic analysis in this work shows that nucleotide homology exists between the 4320 chromosome and that of the pathogenic *Neisseria*. In addition the southern hybridization displayed a strong signal under very stringent conditions.

Because *N. sicca* 4320 diplaying a mutant LPS molecule was not identified by screening mutants with a specific antibody or lectin, directly mutating rhamnose biosyntheis genes, or randomly observing the LPS of hundreds of transformants two explanations are possible. One explanation is that a cluster of genes is responsible for LPS biosynthesis. *Neisseria* are known to direct LOS synthesis from three chromosomal loci, *lgt-1*, *lgt-2* and *lgt-3* (10,48,68). It is possible that a cluster of LPS biosynthetic genes was not cloned into the original plasmid pool if the EcoRI restriction fragment that contained the cluster was too large. Additionally, it is possible that mutations in the LPS biosynthetic genes were contained in the mutant plasmid pool. However, selection of all of the mutants in this study was based on kanamycin resistance. Any alteration in the bacterial membrane can alter the level of resistance to antibiotics when compared to the wild type cells. It is likely that the absence of LPS on the surface of *N. sicca* 4320 dramatically reduces its resistance to kanamycin.

CHAPTER 6. IDENTIFICATION AND CHARACTERIZATION OF *N. GONORRHOEAE* Lpt3

INTRODUCTION

Both the gonococcus and meningococcus decorate the 3, 6, or 7 positions of HepII with PEA (99,101,102). Genetic studies utilizing random transposon mutagenesis of *N. meningitidis* MC58 identified a gene, *lpt3,*that resulted in the expression of LOS devoid of PEA on HepII (81). While enzymatic activity of the *lpt3* gene product has not been verified, transferase activity has been demonstrated in an *Escherichia coli* homolog that possessed the ability to mediate the addition of PEA to the 7-position of KDO (70). Also, the *pmrC* homolog in *Salmonella enterica* was shown to mediate addition of PEA to Lipid A (78). These findings suggested that Lpt3 possesses a corresponding biochemical function (70).

Previous reports, based on Southern hybridization data, suggested that *lpt3* was present in most strains of *N. meningitidis*, but that a homolog was absent in *N. gonorrhoeae* FA1090 (81). However, using the FA1090 DNA sequence database (University of Oklahoma) I identified a homolog, NG1198, that has a 96% nucleotide identity to *N. meningitidis* MC58 *lpt-3*. In this chapter, biochemical evidence that this gene encodes a PEA transferase in the gonococcus is shown.

LOS can be classified based on its ability to bind specific monoclonal antibodies (Mabs). Therefore, the Mabs can be used as a tool to determine the presence of the a particular LOS structure. For example, Fig. 29 is the LOS structure recognized by Mab 2-1-L8. The epitope was contained on a 3.6 kDa component that is commonly expressed by the pathogenic *Neisseria*. (113). Structural analysis on LOS isolated from *N. gonorrhoeae* MS11mk LOS (binds Mab 2-1-L8) defined the carbohydrate portion of the LOS needed

to bind this antibody (Gal β 1-->4Glc β 1-->4[GlcNAc α 1-->2Hep α 1-->3]Hep α 1-->KDO (72). Genetic studies confirmed that Mab 2-1-L8 has affinity for this structure. For example, this Ab binds the LOS resulting when LgtE is expressed, and when LgtA, LgtG, and LgtC are not produced (86). Mab 2-1-L8 loses its affinity for LOS if LgtA, LgtC or LgtG is expressed (18,23,31,109). It was also speculated that the Mab 2-1-L8 specificity requires not only the lactosyl group, but the phosphorylated GlcNAc-Hep residue that is partially exposed on the cell surface (72). This idea was supported as the absence of PEA at 3-HepII from LOS with a composition of $(Hex)_2(Hep)_2(HexNAc)_1(PEA)_1(KDO)_2$ resulted in the loss of Mab 2-1-L8 binding (86). Therefore, Mab 2-1-L8 was used in this study as a tool to screen for gonococcal *lpt3* mutants and to test the biochemical activity of Lpt3.

Since neisserial LOS can modify HepII with PEA at the 6 and 7 positions, it was not known if these modifications effected Mab 2-1-L8 binding to LOS. To clarify this issue, a series of isogenic derivatives expressing LOS with differing PEA modifications was constructed. This allowed for the determination of the effect of PEA modification on Mab 2-1-L8 binding.

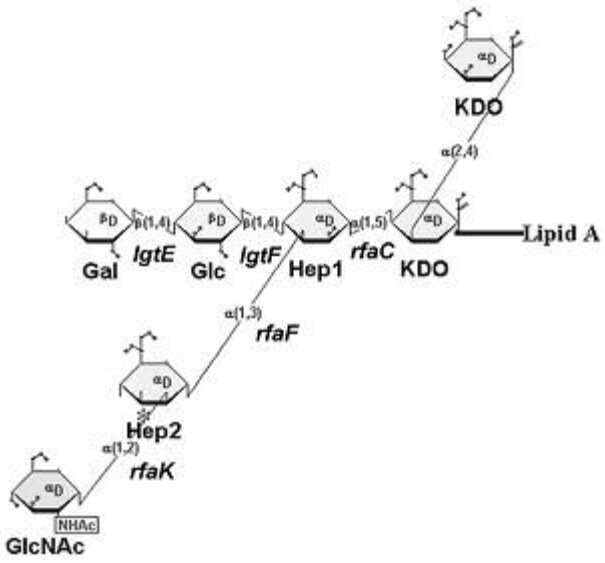


Figure 29. Structure of Neisserial LOS. This structure presented in this figure represents the LOS molecule that can be synthesized by the gonococcus that is the epitope for Mab 2-1-L8 (Modified from Tong *et al.*, (128)). LOS biosynthetic genes are italicized. Phosphoethanolamine (PEA) can is added to the site designated with a * on the HepII residue.

RESULTS

Discovery of *N. gonorrhoeae* FA1090 *lpt3* homolog.

A search of the *N. gonorrhoeae* FA1090 chromosome at the University of Oklahoma with the coding sequence of lpt3 indicated that FA1090 possessed a homolog to that gene. An alignment of the two Lpt3 proteins is shown in Fig. 30. Using the amino acid sequence of Lpt3 to search the NCBI conserved domain website yielded results consistent transferase ability. The region between residues 211 and 497 was 90.6% aligned with pfam00884, a sulfatase. Lpt3 residues 49-464 aligned 71.7% with COG 1368 from a phosphoglycerol transferase of the alkaline phosphatase superfamily. A hydrolase of the alkaline phosphatase superfamily contained COG 3083 which aligned 50.5% with Lpt3. These results indicate that *lpt3* should encode a membrane-associated transferase.

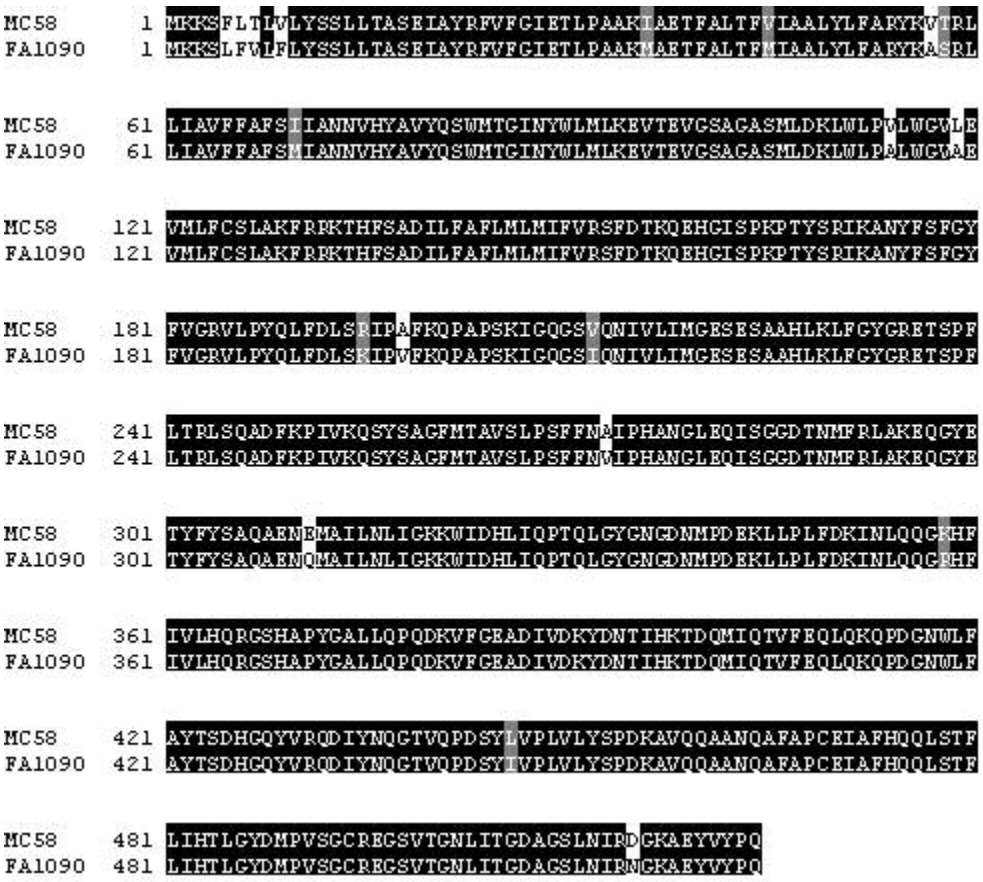


Figure 30. Lpt3 Amino Acid Alignment. *N. meningitidis* MC58 Lpt3 is aligned against *N. gonorrhoeae* FA1090 Lpt3. Darkened regions indicate sites of conservation. Grey boxes indicate where minor amino acid differences exist and white areas show sites of more significant variation.

Screening to find gonococcal LOS transposon mutants.

To determine if Epicente's EZ::Tn system could be used as a mutagenesis tool to knock-out LOS synthesis genes in the gonococcal chromosome, a series of experiments were undertaken to create and identify specific LOS mutants. The strain F62ΔLgtA was chosen as a starting point because it produces an LOS with the structure shown in Fig 29, containing the epitope recognized by Mab 2-1-L8. If the transferase genes previously identified for the production of this structure (*lgtE*, *lgtF*, *rfaK*, *rfaF*, *rfaC*) are inactivated, the resulting LOS structure fails to bind Mab 2-1-L8. Based on these facts, random transposon mutagenesis of F62ΔLgtA chromosomal DNA should result in the inactivation of these genes yielding an altered LOS structure.

F62ΔLgtA chromosomal DNA was extracted and used as a target for random mutagenesis in an *in-vitro* reaction. The resulting DNA was transformed into strain F62ΔLgtA and gonococci were selected based on the acquisition of the kanamycin resistance marker that is carried within the transposon. To verify randomness chromosomal DNA was isolated from individual transformants and used in the Southern hybridization shown in Fig 31. Hybridization to fragments of different mobilities is evidence that transposon insertions into the chromosome occurred in a random pattern.

The transformants were patch plated and screened for reactivity with Mab2-1-L8. An example of one of these colony blots is shown in Fig. 32. This blot clearly shows that two types of transformants existed, one group that reacted with the Mab and one group that did not. Mutants that reacted with the Mab 2-1-L8 likely contained a transposon insertion into a gene that did not affect the production of LOS. The non-reactive

transformants were selected for further study as it was possible that produced LOS molecules had a structure varying from F62ΔLgtA LOS.

LOS was prepared from each of the Mab2-1-L8 non-reactive transformants and analyzed by SDS-PAGE. As shown in Fig. 31 the LOS molecules from the non-reactive mutants had various degrees of mobilites. Some mutant LOS molecules, such as row 1 lane 6, have increased mobility compared to the F62ΔLgtA parent. This truncation of LOS is best explained by the inactivation of a transferase enzyme that is necessary to produce the parent structure. These results are consistent with the transposon insertions affecting the LOS type of the mutants that show increased mobility. Also, mutant LOS molecules were seen that had an increased mobility, such as in row 1 lane 7. This result was likely due to phase variation within *lgtC* or *lgtG* that allowed these genes to produce functional proteins to respectively add residues to the α and γ chains (see Fig. 1).

To determine if the transposon insertions were responsible for the truncated LOS mutants, the sites of transposon insertion were established. Chromosomal DNA was extracted from the LOS mutants and each separately digested with EcoRI. The digestion reaction was diluted and chromosomal fragments were ligated on themselves. The mixture of ligated fragments was used as a template for PCR with primers Kan-FP and Kan-RP that bind to the ends of the transposon and are oriented to so that amplification of the sequence flanking the transposon can occur. Sequencing of the PCR products showed that the transposon had inserted into LOS transferase genes such as *lgtE* and *rfkA*, the LOS mobilities of these mutants are respectively shown in Fig. 33 panel A lanes 9 and 10. These results show that Epicentre's EZ::Tn technology can be used to create gene

knock-outs in the gonococcus and that these insertions can alter the produced LOS molecule. Therefore, the transposon system should be able to inactivate gonococcal *lpt3*.

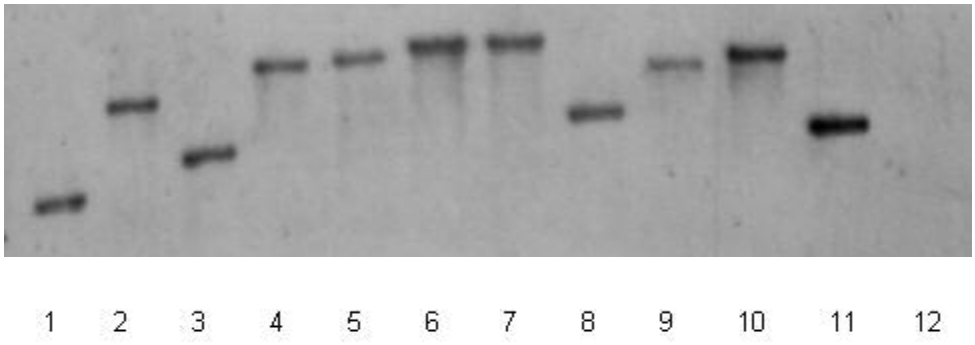


Figure 31. Southern blot of random F62ΔLgtA transposon mutants. Chromosomes from kanamycin resistant F62ΔLgtA mutants were extracted and digested with EcoRI and subjected to hybridization with the Kan^r marker from Tn5. Lanes 1 through 11 contain the random transposon mutant chromosomes and lane 12 contains the F62ΔLgtA parent chromosome.

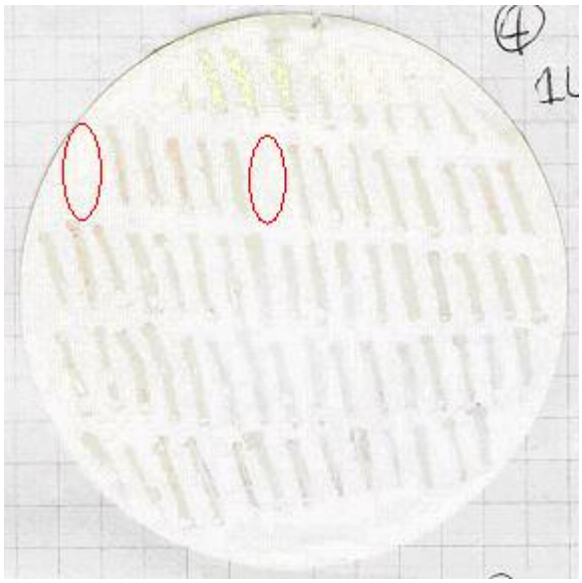


Figure 32. Colony blot of *N. gonorrhoeae* F62ΔLgtA transposon mutants with mAb 2-1-L8. Each colony blot screened approximately 70 mutants for binding affinity. In this blot 2 of the transformants (Row 2 streaks 1 and 7) failed to react with Mab 2-1-L8.

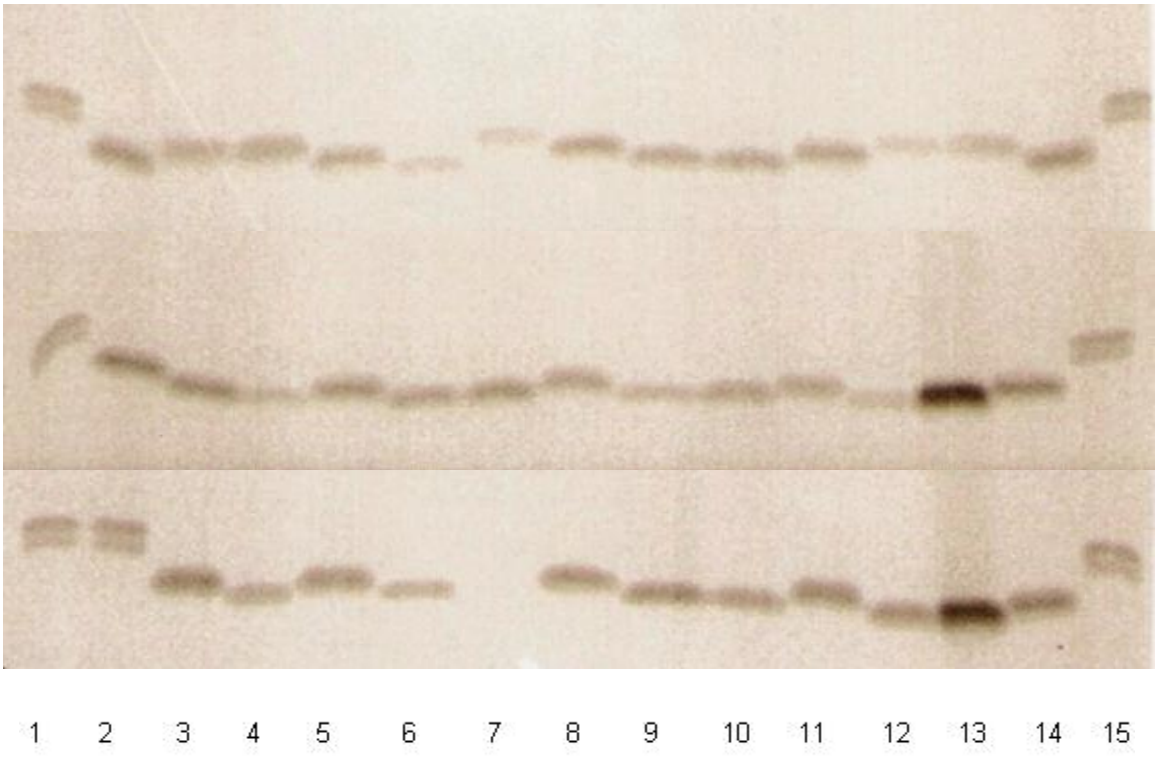


Figure 33. SDS-PAGE of LOS from *N. gonorrhoeae* F62ΔLgtA mAb 2-1-L8 non-reactive transposon mutants. Shown are 3 silver stained gels with the collection of all of the non-reactive mutant. Lanes on the end (1 and 15) showing two bands contain F62 LOS. Lanes 2, 5, 8, 11, and 14 contain F62ΔLgtA LOS. This parent sample is in every third lane. Two transposon mutant LOS samples are shown between the parent samples.

Construction of gonococcal Lpt3 mutant.

In order to study the activity of the *lpt3* gene product, *N. gonorrhoeae* F62ΔLgtA was chosen as a starting point because F62ΔLgtA contains a deletion of the *lgtA* gene, resulting in the expression of a single LOS component that binds the Mab 2-1-L8 (see Fig. 29). As shown in Fig 34, using primers PTE-4 and PEA-R, a PCR was performed using the F62ΔLgtA chromosome as a template. The DNA sequence of the resulting 1.8 kb amplicon was determined to contain *lpt3* and as well as a neisserial uptake sequence outside of the coding region. The region encoding *lpt3* showed 100% nucleotide identity to *N. gonorrhoeae* FA1090. The flanking regions contained only a few base changes in non-coding regions (data not shown). This fragment was cleaved with restriction enzymes EcoRI and NdeI and cloned into pUC19 at those respective sites. The resulting plasmid, pLPT3, was used as a target of transposon mutagenesis.

Transposon inserted pLPT3 was used to transform F62ΔLgtA, and kanamycin resistant transformants identified. These colonies were screened for loss of reactivity with Mab 2-1-L8. PCRs were performed on individual Mab 2-1-L8 unreactive transformants using PTE-4 and PTE-5 and the amplicons all had the predicted increase in size, corresponding to the transposon insertion into *lpt3*. DNA sequence analysis of the amplicons confirmed that each contained a transposon insertion into *lpt3*. One of these mutants, F62ΔLgtAlpt3::Tn5 was selected for further study. The data presented in Fig 35 indicates that the SDS-PAGE profile of the mutant strain produced a LOS that migrated slightly faster than F62ΔLgtA LOS. Western blot analysis of an identical gel, using Mab 2-1-L8 indicated that this component no longer bound the Mab. Because the antibody

requires PEA addition at 3-HepII, this mutant presumably contained an insertion that inactivated *lpt3* and prevented this specific decoration.

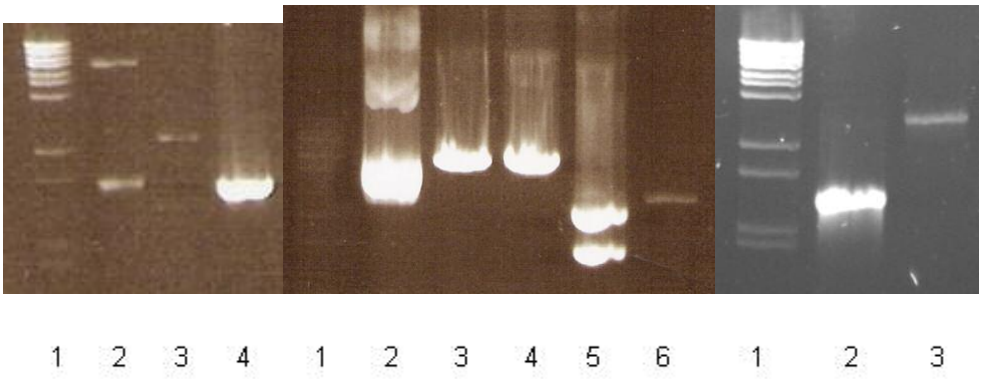


Figure 34. Construction of *N. gonorrhoeae* F62ΔLgtAlpt3::Tn5. Lanes 1-4 in Panel A (left) respectively include λ BstEII standard, pUC19, pUC19 EcoRI and NdeI, and *lpt3* amplification product cleaved with EcoRI and NdeI. Lanes 1-6 in Panel B (middle) respectively contain λ BstEII standard, pLPT3, pLPT3 EcoRI, pLPT3 NdeI, pLPT3 EcoRI and NdeI, and pUC19 EcoRI. Panel C (right) shows the *lpt3* PCR product from F62ΔLgtA in lane 2 compared to the *lpt3* PCR product from F62ΔLgtAlpt3::Tn5.

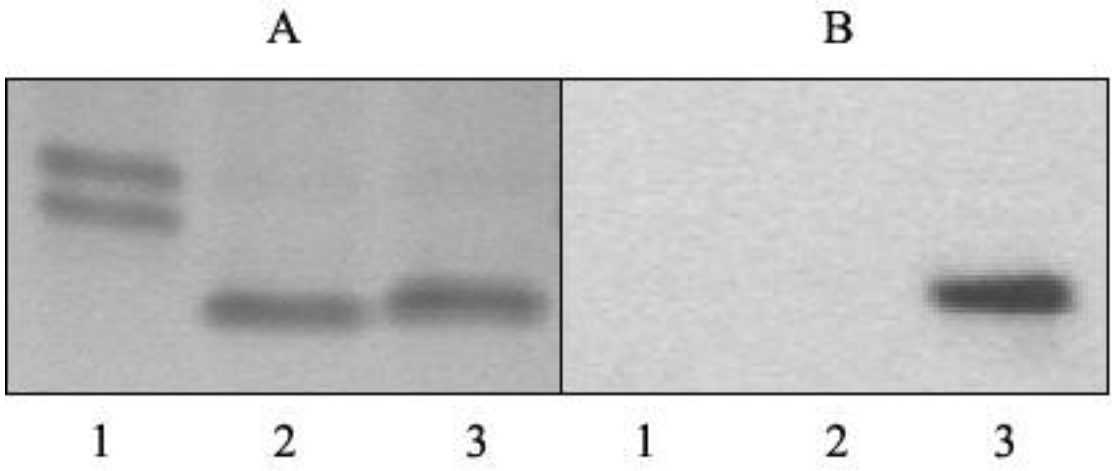


Figure 35. SDS-PAGE and Western blot of *N. gonorrhoeae* F62ΔLgtAlpt3::Tn5 LOS with mAb 2-1-L8. Panel A shows silver stained LOS electrophoresed through a 16.5% Tris-Tricine polyacrylamide gel. The lanes represent: 1) LOS isolated from F62; 2) LOS isolated from F62ΔLgtAlpt3::Tn5; and 3) LOS isolated from F62ΔLgtA. Panel B represents a western blot on an identical gel as shown in panel A, using Mab 2-1-L8.

Complementation analysis of gonococcal Lpt3 mutant.

To demonstrate that the loss of Mab 2-1-L8 affinity for F62ΔLgtAlpt3::Tn5 LOS was caused by the inactivation of *lpt3,* complementation analysis was performed. The plasmid pLPT3, containing wild type gonococcal *lpt3*, was used as a source of DNA to transform F62ΔLgtAlpt3::Tn5 cells; kanamycin sensitive transformants were identified for further study. PCR analysis with primers PTE-4 and PTE-5 in Fig. 36 revealed that all kanamycin sensitive transformants lost the transposon insertion, and contained an amplicon of the same size as the original parent. All kanamycin sensitive transformants reacquired the ability to bind Mab 2-1-L8. One of these transformants was selected for further study. LOS was isolated from this complemented mutant and its SDS-PAGE profile compared to that of F62ΔLgtA and F62ΔLgtAlpt3::Tn5. The data indicated (Fig. 37) that the complemented mutant and the parent showed identical LOS migration patterns. Western blot analysis of a duplicate gel indicate that both the parent and the complemented mutant LOS bound the mAb, while the *lpt3* mutant LOS did not. These results indicate that the insertional inactivation of *lpt3* was responsible for the inability of mutant LOS to bind Mab2-1-L8.

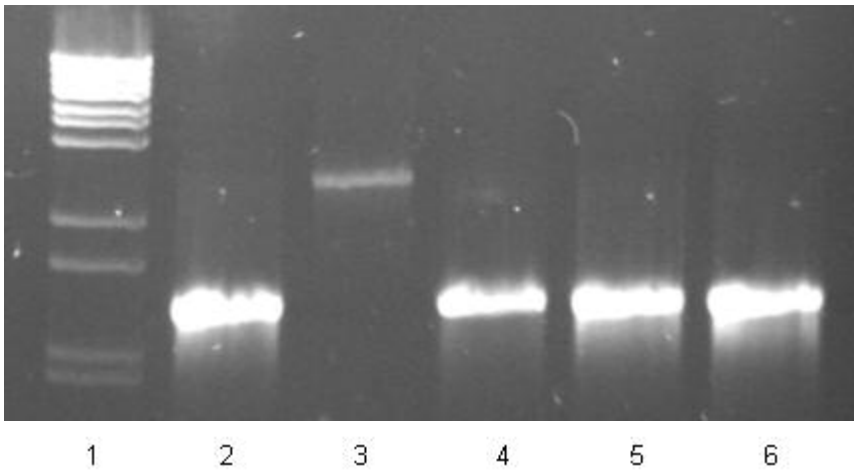


Figure 36. Colony PCRs of *N. gonorrhoeae* F62ΔLgtA lpt3::Tn5 complemented mutants. Lane 1 shows the λ BstEII standard. Lanes 2 and 3 respectively contain the *lpt3* PCR products from F62ΔLgtA and F62ΔLgtAlpt3::Tn5. Lanes 4-6 show the *lpt3* PCR product for F62ΔLgtAlpt3::Tn5 complemented with pLPT3.

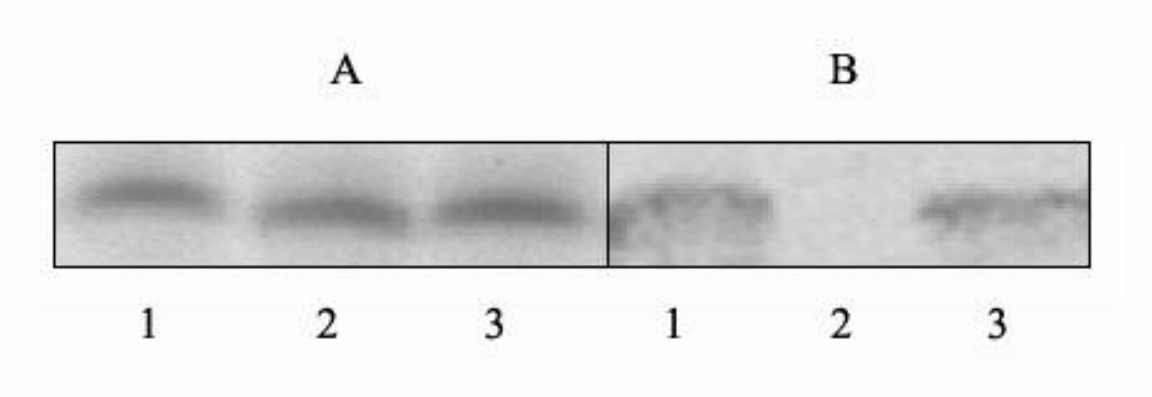


Figure 37. SDS-PAGE and Western blot with mAb 2-1-L8 of *N. gonorrhoeae* F62ΔLgtA lpt3::Tn5 complemented mutant. Panel A shows silver stained LOS electrophoresed through a 16.5% Tris-Tricene polyacrylamide gel. The lanes show: 1) LOS isolated from F62ΔLgtA; 2) LOS isolated from F62ΔLgtAlpt3::Tn5; and 3) LOS isolated from the *lpt3* complemented F62ΔLgtAlpt3::Tn5. Panel B depicts a western blot on an identical gel as shown in panel A, using Mab 2-1-L8.

Construction of gonococcal lpt6 an lpt3lpt6 mutants.

N. gonorrhoeae can decorate HepII with PEA at the 3, 6, and 7 positions. In order to show that the loss of Mab 2-1-L8 reactivity is caused by the lack of PEA at 3-HepII, a series of F62ΔLgtA strains were engineered to express LOS molecules that were devoid of PEA addition at varying sites on HepII. *N. gonorrhoeae* F62ΔLgtA was selected as the parent strain for this study because it expresses the LOS structure shown in Fig. 29 which is recognized by Mab 2-1-L8. LOS devoid of PEA at 3-HepII was expressed by F62ΔLgtAlpt3::Tn5, as Lpt3 is shown in this study to act as a PEA transferase at that site. To obtain LOS devoid of PEA at 6-HepII, *lpt6* was deleted from the chromosome of *N. gonorrhoeae* F62ΔLgtA (*lpt6* is speculated to encode a PEA transferase that acts at that site (139). *N. gonorrhoeae lpt6* was cloned by PCR amplification of *lpt6* its flanking regions with primers O6-PEAR and O6-PEAF. The amplicon was digested with EcoRI and HindIII and cloned into the EcoRI and HindIII sites of pUC19, giving pLPT6. The *lpt6* gene was replaced with a $Spec^r$ cassette by using pLPT6 as a template in a PCR reaction with mutation primers Lpt6 MutF and Lpt6 MutR The amplicon was digested with PstI, the site into which the $Spec^r$ cassette was inserted. This construct was introduced into F62ΔLgtA, giving rise to F62ΔLgtAΔlpt6. A third strain was constructed which failed to express Lpt3 and Lpt6. This strain, F62ΔLgtAΔlpt6lpt3::Tn5 was created by transforming *N. gonorrhoeae* F62ΔLgtAlpt3::Tn5 cells with the *lpt6* deletion construct.

MALDI-MS Analysis of Purified Mutant LOS.

Mass spectrometry analysis was performed to confirm the absence of PEA in LOSs expressed by strains containing mutant *lpt3* and *lpt6*. The spectras of LOSs isolated from strains F62ΔLgtA, F62ΔLgtAlpt3::Tn5, lpt3 complemented mutant, F62ΔLgtAΔlpt6, and F62ΔLgtAΔlpt6lpt3::Tn5 are shown in Fig 38. The proposed structure of molecules corresponding to each major ion are listed in Table 7. As shown in the spectra in Fig. 38 panel A, the most abundant ion for F62ΔLgtA is at *m/z* 2550.48. This value is consistent with the composition $(Hex)_2(HexNAc)_1(Hep)_2(PEA)_2(KDO)_2$ and a lipoidal moiety (952 Da). The presence of minor peaks in the spectra allowed us to obtain further structural validation. The peak representing the *m/z* 2427.19 ion differs from the major peak by the mass of PEA (123.1 Da). The *m/z* 2330.64 ion corresponds to the loss of KDO (220.2 Da) when compared to the major ion. This molecule occurs as the result of the lability of the ketosidic linkage resulting from hydrazine mediated de-O-acetylation (64). When compared to the most abundant ion, the *m/z* 2630.07 ion differs by the gain of HexNAc (203.2 Da) and the loss of KDO and PEA. A mass difference consistent with the gain of a Hex (162.2 Da) is seen in the *m/z* 2713.04 molecular ion when compared to the *m/z* 2550.48 ion.

Comparison of the F62ΔLgtA and F62ΔLgtAlpt3::Tn5 LOS MALDI-MS spectra showed compositional differences in the LOS produced by these two strains. The spectra in panel B shows that the major ion of the *lpt3* mutant LOS is at *m/z* 2426.69, corresponding to a loss of PEA when compared to the most abundant peak in the parental spectrum. Ions of *m/z* 2329.09 and 2206.69 are consistent with the loss of KDO from ions *m/z* 2426.69 and 2550.00, respectively. Both the parent and mutant spectra contain a peak

at m/z 952 representing the lipoidal moiety. F62ΔLgtA does not decorate its lipoidal moiety with appreciable PEA residues, as shown by the absence of peaks at 1032, 1155, 1075 and 1278, which would correspond to the addition of one or two phosphate groups and one or two PEA groups respectively to the lipoidal moiety. Therefore, loss of PEA groups in the family of PEA transferase mutants created from this strain must be from the oligosaccharide component of the LOS.

LOS from the complemented *lpt3* mutant LOS was also analyzed by mass spectrometry. The spectra in panel C indicates that there was a recovery of wild type LOS. This is best viewed through the change in ratios of ions containing PEA decoration in comparison to the mutant spectra. The major peak is *m/z* 2550.48, characteristic of $(Hex)_2(HexNAc)_1(Hep)_2(PEA)_2(KDO)_2$ and the lipoidal moiety. Again, like the wild type spectra, a minor ion of *m/z* 2427.19 is seen which differs from the major ion by a PEA residue. Other similarities to the wild type spectra are three ions corresponding to the loss of KDO, a gain of HexNAc with a simultaneous loss of KDO and PEA, and the gain of a Hex residue. These results indicate that inactivation of Lpt3 is responsible for the loss of PEA shown in the mutant spectra.

The MALDI-MS spectra of F62ΔLgtAΔlpt6 LOS also showed compositional differences when compared to the parent molecule. The spectra in panel E shows that the major ion of the *lpt6* mutant LOS was 2426.05 m/z which is consistent with $(Hex)_2(HexNAc)_1(Hep)_2(PEA)_1(KDO)_2$ lipoidal moiety and the loss of a PEA residue. The presence of the 2207.65 m/z ion is consistent with the loss of a KDO, due to the lability of the ketosidic linkage, as well as the loss of a PEA residue.

Both Lpt3 and Lpt6 mutant LOSs contained small amounts of ions that contain two PEA residues. For example, the Lpt3 mutant spectra shows an ion at *m/z* 2550.00 corresponding to $(Hex)_2(HexNAc)_1(Hep)_2(PEA)_2(KDO)_2$ lipoidal moiety. Also, the Lpt6 mutant spectra contains a peak at 2329 m/z corresponding to $(Hex)_2(HexNAc)_1(Hep)_2(PEA)_2(KDO)_1$ lipoidal moiety. The presence of these ions is explained by promiscuity of the PEA transferases. Lpt6, which is responsible for the addition of exocyclic PEA, can also function to decorate 3-HepII. Lpt3, which is responsible for the addition of cyclic PEA, can also function to decorate 6-HepII.

MALDI-MS was also performed on LOS expressed by F62ΔLgtAΔlpt6lpt3::Tn5. The results of this analysis are shown in panel D. This LOS is completely devoid of PEA decoration as demonstrated by the *m/z* 2304.51 ion corresponding to the removal of both PEA groups when compared to the major ion of *m/z* 2550.48 in the parent spectra. This data indicates that Lpt6 is responsible for the two PEA modifications in the *lpt3* mutant and that Lpt3 is responsible for the two PEA additions in the *lpt6* mutant. The Lpt3 and Lpt6 PEA transferases are likely promiscuous as to their sites of action, albeit with different affinities for the two sites of adornment on HepII, cyclic and exocyclic. Lpt6 appears to prefer the exocyclic sites, 6-HepII and 7-HepII, whereas Lpt3 appears to have a higher affinity 3-HepII than the exocyclic sites.

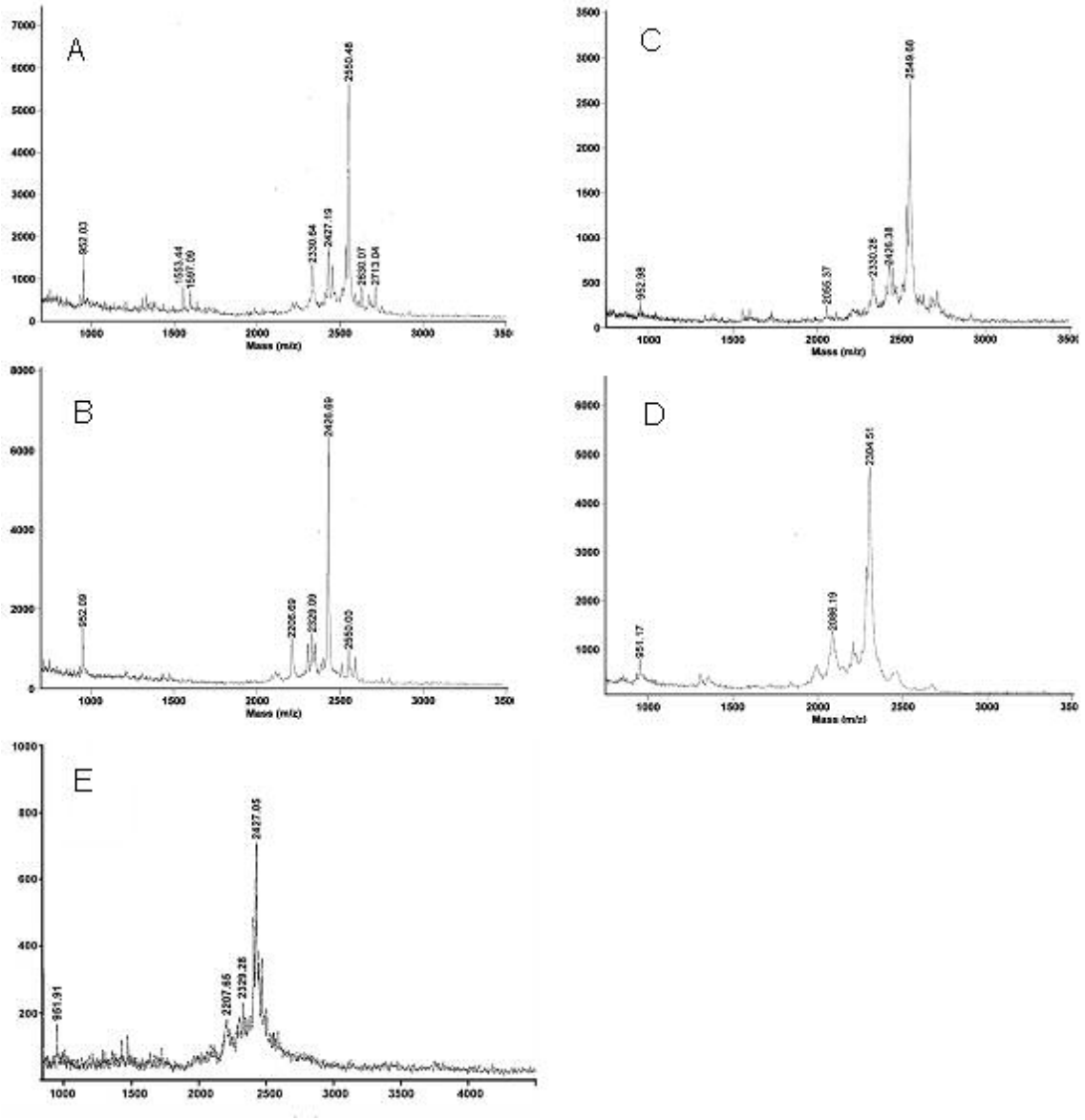


Figure 38. MALDI-TOF analysis of *N. gonorrhoeae* F62 PEA Mutant LOS. Panels A, B, C, D, and E show spectra of O-deacylated purified LOS. Panel A shows the mass spectrum profile of F62ΔlgtA LOS. Panel B shows the spectrum produced by F62ΔLgtAlpt3::Tn5. Panel C shows the spectrum produced by the complemented mutant. Panel D shows the F62ΔLgtAΔlpt6lpt3::Tn5 LOS profile. Panel E shows the spectrum produced by F62ΔLgtAΔlpt6 LOS. The masses of the abundant fragments are indicated at the top of the corresponding peaks.

Table 7. Compositions and masses of LOS from MALDI-MS

Strain	LOS Composition[a]	Experimental M_r	Calculated M_r
F62ΔLgtA			
	$(Hex)_2(HexNAc)_1(Hep)_2(PEA)_2(KDO)_2$ lipoidal moiety	2550.48	2551.27
	$(Hex)_2(HexNAc)_1(Hep)_2(PEA)_1(KDO)_2$ lipoidal moiety	2427.19	2428.22
	$(Hex)_2(HexNAc)_1(Hep)_2(PEA)_2(KDO)_1$ lipoidal moiety	2330.64	2331.09
	$(Hex)_2(HexNAc)_2(Hep)_2(PEA)_1(KDO)_1$ lipoidal moiety	2630.07	2631.41
	$(Hex)_3(HexNAc)_1(Hep)_2(PEA)_2(KDO)_2$ lipoidal moiety	2713.04	2713.83
F62ΔLgtAlpt3::Tn5			
	$(Hex)_2(HexNAc)_1(Hep)_2(PEA)_1(KDO)_2$ lipoidal moiety	2426.69	2428.22
	$(Hex)_2(HexNAc)_1(Hep)_2(PEA)_2(KDO)_1$ lipoidal moiety	2329.09	2331.09
	$(Hex)_2(HexNAc)_1(Hep)_2(PEA)_1(KDO)_1$ lipoidal moiety	2206.69	2208.04
	$(Hex)_2(HexNAc)_1(Hep)_2(PEA)_2(KDO)_2$ lipoidal moiety	2550.00	2551.27
F62ΔLgtAlpt3::Tn5 pLPT3 complement			
	$(Hex)_2(HexNAc)_1(Hep)_2(PEA)_2(KDO)_2$ lipoidal moiety	2549.68	2551.27
	$(Hex)_2(HexNAc)_1(Hep)_2(PEA)_1(KDO)_2$ lipoidal moiety	2426.38	2428.22
	$(Hex)_2(HexNAc)_1(Hep)_2(PEA)_2(KDO)_1$ lipoidal moiety	2330.28	2331.28
	$(Hex)_2(HexNAc)_2(Hep)_2(PEA)_1(KDO)_1$ lipoidal moiety	2629.99	2631.41
	$(Hex)_3(HexNAc)_1(Hep)_2(PEA)_2(KDO)_2$ lipoidal moiety	2712.96	2713.83
F62ΔLgtAΔlpt6			
	$(Hex)_2(HexNAc)_1(Hep)_2(PEA)_1(KDO)_2$ lipoidal moiety	2427.05	2428.22
	$(Hex)_2(HexNAc)_1(Hep)_2(PEA)_2(KDO)_1$ lipoidal moiety	2329.28	2331.28
	$(Hex)_2(HexNAc)_1(Hep)_2(PEA)_1(KDO)_1$ lipoidal moiety	2207.65	2208.04
F62ΔLgtAΔlpt6lpt3::Tn5			
	$(Hex)_2(HexNAc)_1(Hep)_2(PEA)_0(KDO)_2$ lipoidal moiety	2304.51	2305.17
	$(Hex)_2(HexNAc)_1(Hep)_2(PEA)_0(KDO)_1$ lipoidal moiety	2086.19	2084.99

Structures given are the predicted structure. The calculated M_r is derived from this structure.

Determination of Mab 2-1-L8 Specificity for Mutant LOS.

To visualize the PEA mutant molecules, LOS from strains F62, F62ΔLgtA, F62ΔLgtAlpt3::Tn5, F62ΔLgtAΔlpt6, and F62ΔLgtAΔlpt6lpt3::Tn5 were subjected to SDS-PAGE. The electrophoretic mobilities of each of these preparations are shown in Fig. 39. In comparison to F62ΔLgtA LOS the molecules produced by F62ΔLgtAlpt3::Tn5, F62ΔLgtAΔlpt6, and F62ΔLgtAΔlpt6lpt3::Tn5 exhibited similar rates of migration. PEA residues contain a positive charge, and the loss of charge retards the LOS mobility to the same degree that the loss of mass increases it. Since the MALDI-MS results showed that PEA residues had been removed from the F62ΔLgtA LOS, the specificity of Mab 2-1-L8 could be determined for each mutant LOS. To test the affinities, a Western blot was performed with Mab 2-1-L8 on an identical gel. After SDS-PAGE, LOSs were electrotransferred onto a membrane which was processed in blocking buffer and screened for reactivity with Mab 2-1-L8. The results of the Western blot are shown in Fig 3b. As expected, F62 LOS failed to react with Mab 2-1-L8 while its truncated LOS created by the *lgtA* deletion recognized the antibody. LOS produced by F62ΔLgtAlpt3::Tn5 that was devoid of PEA at 3-HepII failed to bind Mab 2-1-L8. This is consistent with the report that F62ΔLgtA LOS molecules must have a PEA substituent on 3-HepII in order for them to bind Mab 2-1-L8. The blot also showed that LOS produced by F62ΔLgtAΔlpt6 that lacked PEA decoration at 6-HepII reacted with Mab 2-1-L8 while the LOS that was devoid of modification at both the 3 and 6 positions of HepII failed to bind the antibody. These results show that an exocyclic PEA substitution (6 or 7-HepII) is not adequate to shape the LOS structure that binds Mab 2-1-L8, in the absence of a PEA at 3-HepII.

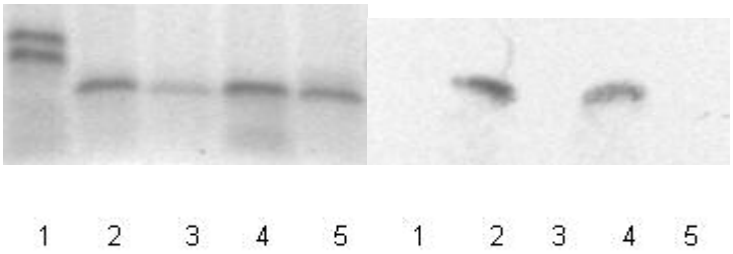


Figure 39. SDS-PAGE and Western blot with Mab 2-1-L8 of *N. gonorrhoeae* family of PEA mutants. Lanes 1-5 in Panel A respectively show LOS molecules extracted from F62, F62ΔLgtA, F62ΔLgtAlpt3::Tn5, F62ΔLgtAΔlpt6, F62ΔLgtAΔlpt6lpt3::Tn5. Panel B shows the Western blot of an identical gel with Mab 2-1-L8.

Biochemical analysis of Lpt3.

In order to demonstrate that the *lpt3* gene product was a PEA transferase, Lpt3 was isolated and purified in conjunction with A. Piekarowicz. Two PCR reactions were performed using FA1090 chromosomal DNA as a template, one using primers PTE-6 and PTE-7 and the other with primers PTE-4 and PTE-5. Each product yielded a 1.5 kb fragment that were both confirmed by nucleotide sequencing to contain *lpt3*. Each amplicon was cloned respectively into expression vectors pQE30 and pET15b, creating pQE30::PEA and pET15b::PEA. These plasmid DNAs were transformed into *E. coli* M15 (for vector pQE30) or BL21 (for vector pET15b). Expression in these vectors creates a fusion protein in which a 6 residue His tag is added to Lpt3. As shown in Fig. 40, the protein fractions taken immediately after the end of induction of BL21 cells containing pET15b::PEA showed the presence of a protein of the appropriate molecular mass (54 Kda). After purification on a nickel column, even in the presence of the protease inhibitor PMSF, a degradation product of a lower molecular mass was always observed.

The activity of Lpt3 was determined using whole cells of F62ΔLgtAlpt3::Tn5. LOS was isolated from the F62ΔLgtAlpt3::Tn5, with and without enzymatic treatment and subjected to SDS-PAGE and western blot analysis. As shown in Fig. 41, silver staining of the gel revealed that, as expected, the F62ΔLgtAlpt3::Tn5 cells that underwent the reaction with or without the enzyme yielded a slightly different migration pattern. LOS that underwent reaction with Lpt3 migrated at a reduced rate compared to the mutant LOS. The difference is difficult to see because the gel contained overloaded LOS which was required for western blotting. The Western blot of an identical gel showed a

difference in Mab 2-1-L8 reactivity. Only F62ΔLgtAlpt3::Tn5 cells that underwent the reaction with the Lpt3 enzyme regained the ability to bind Mab 2-1-L8. This suggests that PEA had been transferred to the 3-HepII on the gonococcal LOS.

Because LOS was contained in whole cells in the biochemical experiment described above, it is possible that some additional components contained in these cells were needed for the successful LOS modification. As a direct measure of Lpt3 modification of LOS, LOS from F62ΔLgtAlpt3::Tn5 was isolated and purified and an enzymatic reaction was performed. As shown in Figure 42, purified mutant LOS incubated with phosphatidylethanolamine and the purified Lpt3 enzyme produced a strong positive reaction with Mab 2-1-L8, when the reaction mix was directly spotted onto a nitrocellulose filter. When the reaction was performed either without PEA or without Lpt3, no Mab 2-1-L8 reactivity was seen. As an additional control, lysate from sonicated *E. coli* BL21 cells were reacted with LOS from F62ΔLgtAlpt3::Tn5. Reactivity with Mab 2-1-L8 was never seen (data not shown). This control eliminates the possibility that gonococcal Lpt3 has been contaminated with a PEA transferase from the cloning host. Attempts to demonstrate this reactivity by SDS-PAGE analysis failed to produce a satisfactory gel. This is most likely due to the fact that we needed to add large quantities of LOS (3 μg) to get a readily visible signal when the modified LOS was blotted directly onto nitrocellulose, and this quantity of LOS grossly overloads and SDS-PAGE gel. However, the ability of purified Lpt3 to add PEA to purified LOS provides direct evidence that *lpt3* encodes a functional PEA transferase that adds PEA onto 3-HepII of gonococcal LOS.

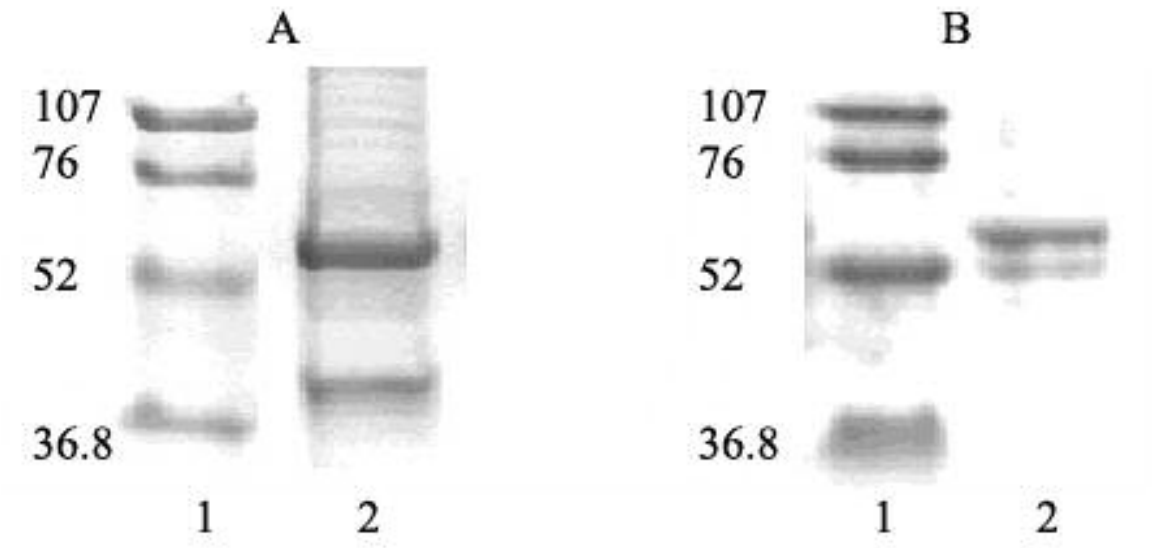


Figure 40. Isolation of *N. gonorrhoeae* FA1090 Lpt3. *E. coli* strains were grown to mid-log phase, and Lpt3 expression was induced by the addition of IPTG. In Panel A, cell supernatants were analyzed on a 16.5% SDS-PAGE gel. Lane 1 shows the molecular weight standard and lane 2 shows the extracted supernatants. Panel B shows the Lpt3 protein that was purified on a nickel column. Lane 1 shows the molecular weight standard and lane 2 shows the purified Lpt3. A degradation product was obtained that did not show enzymatic activity in subsequent steps (data not shown). The molecular mass standard is given in Kda.

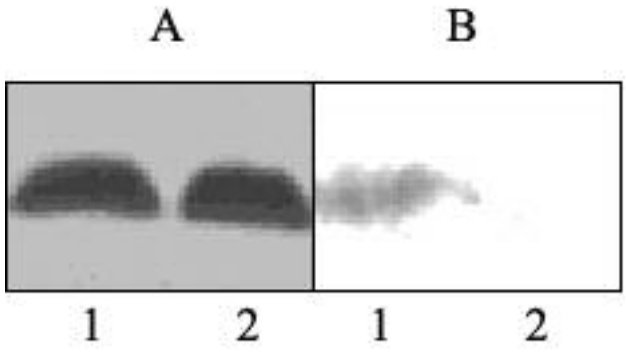


Figure 41. Enzymatic Lpt3 activity on whole F62ΔLgtAlpt3::Tn5 cells. Reactions contained 10 µl of mutant cell suspension, 500 mM Tris-HCl buffer (pH 8.8), 2 µl of Phosphatidylethanolamine, 3 µl of 100mM $MgCl_2$, 2 µl of Lpt3 enzyme, and water to a final volume of 30 µl. The reactions were carried out for 90 minutes at 30^{o}C. 10 µl of the reaction sample was withdrawn and samples were analyzed on a 15 % Tris-Tricine gel. Panel A shows a silver stained gel of LOS resulting from the enzymatic reactions. Lane 1 contains F62ΔLgtAlpt3::Tn5 LOS from a reaction containing Lpt3. Lane 2 shows the LOS resulting from the same reaction, but in the absence of Lpt3. Panel B shows a Western Blot with Mab 2-1-L8 of a gel identical to the gel shown in panel A.

Figure 42. Enzymatic Lpt3 activity on purified LOS. Purified LOS from F62ΔLgtAlpt3::Tn5 was reacted with purified Lpt3. Reactions contained 3 µl of mutant LOS, 500 mM Tris-HCl buffer (pH 8.8), 2 µl of phosphatidylethanolamine, 3 µl of 100mM $MgCl_2$, 2 µl of Lpt3 enzyme, and water to a final volume of 30 µl. The reactions were carried out for 90 minutes at 30^{o}C. 10 µl of each reaction was spotted on the nitrocellulose filter. Lane 1 represents the above reaction, lane 2 shows the complete reaction without Lpt3, and lane 3 shows the complete reaction without phosphatidylethanolamine.

DISCUSSION

This study used bioinformatic and genetic approaches to identify a *lpt3* homolog in *N. gonorrhoeae*. I constructed a mutation in the identified DNA sequence by *in vitro* transposon mutagenesis, introduced the mutated DNA sequence into the gonococcus and analyzed transformants for their ability to bind Mab 2-1-L8. Failure of mutant LOS to bind Mab 2-1-L8 in western blots and mass spectral analyses of this mutant LOS confirmed that the inactivation of gonococcal *lpt3* resulted in the production of LOS that was deficient in PEA modification at 3-HepII. Genetic complementation by restoring the parental DNA sequence into transposon-mutagenized strains restored the parental LOS phenotype, verifying that the loss of reactivity was due to the insertion of the transposon into *lpt3*. While it is possible that the loss of Mab binding could have been due to a polar effect of the transposon insertion effecting the expression of downstream genes, this possibility has been discounted because DNA sequence analysis of both *N. gonorrhoeae* F62 and FA1090 strains revealed that the downstream gene (*hrpA* homolog) has significant homology with DNA helicases and that the intergenic region between *lpt3* and the *hrpA* homolog contain significant homology with promoter elements. In addition, purified Lpt3 was able to mediate the proposed biochemical function of *lpt3*.

A cloning/purification procedure that allowed for purification of Lpt3 was devised. This protein proved difficult to purify to homogeneity, and rapidly degraded in solution. Amino acid sequence analysis of this protein suggests that it is intimately associated with the bacterial membrane. Biochemical tests using the purified protein clearly indicate that Lpt3 is a PEA transferase because it modified 3-HepII of *lpt3* mutant LOS. However, it does not function well in solution. The low activity, and the high

degree of instability of the purified protein, prevented a more detailed biochemical analysis of Lpt3.

Both *N. gonorrhoeae* and *N. meningitidis* are capable of expressing LOS molecules that are decorated with PEA at the 3, 6, and 7 positions of HepII. The exact locations of the addition(s) are biologically significant because Neisserial LOS has been shown to be a target of the immune system. For example, antibodies recognizing the B5 epitope, PEA bound to 3-HepII, were found in patients with meningococcal meningitidis (97). Also, the position of PEA addition has been shown to be important in the role LOS plays as an acceptor for complement. PEA present at 6-HepII residue was shown to form amide linkages with C4b better than corresponding LOS structures with PEA added to the 3-position. As a result, meningococcal strains containing PEA at 6-HepII are more susceptible to the bactericidial effects of serum than strains with the 3-position modification (102). It has been suggested that this difference in C4b interaction depending on the PEA modification site may account for that fact that over 70% of meningococcal strains isolated from humans contain PEA at 3-HepII (98) and why during an outbreak of meningococcal meningitis, 97% of disease causing isolates carried the same modification (65).

A study of Norwegian meningococcal isolates found a significant occurrence of LOS recognized by Mab 2-1-L8 in the population, showing that this epitope is critical to vaccine creation (3). The results of this study show that the site of PEA decoration is crucial to the formation of the epitope as cyclic PEA is required for Mab 2-1-L8 recognition and exocyclic PEA is dispensable. The conformation of the LOS expressed on the surface is partially determined by the electrostatic bonds that form between the

PEA residues and the carboxyl groups of the KDO residues. Therefore, the location of the PEA residues can influence this conformation. PEA at the exocyclic position brings the NH_2 residue closer to the KDO carboxyl residues than PEA at the cyclic position. The stereochemistry of the LOS needs to be studied in order to understans their full immunogen potential.

A closer examination of the MALDI-MS results provides insight into the action of PEA transferases. The F62ΔLgtAlpt3::Tn5 LOS spectrum contained peaks at *m/z* 2426.69 and 2550.00 indicating that these ions respectively contained one and two PEA residues, even though *lpt3* had been inactivated. The Western blot with Mab 2-1-L8 in this study shows that *lpt3* inactivation, results in the loss of PEA decoration at 3-HepII. However, the *lpt3* mutant produced small amounts of LOS decorated with PEA at sites other than 3-HepII. The production of LOS devoid of PEA by F62ΔLgtAΔlpt6lpt3::Tn5 indicates that Lpt6 is responsible for the additional decorations at sites other than 3-HepII in the *lpt3* mutant. LOS produced by the *lpt6* mutant also included small amounts of molecules decorated with two PEA residues. This indicates that the PEA additions in the *lpt6* mutant were made by Lpt3. The genetic and biochemical basis for the addition of PEA onto the 7 position is unknown. However, the inactivation of both *lpt3* and *lpt6* result in the production of LOS that is completely devoid of PEA. One possibility is that PEA at 6-HepII of LOS may spontaneously migrate to the 7-HepII position. In mutants that cannot add PEA onto this site, no addition would be possible at position 7. A second possibility is that the protein needed to mediate this addition requires the presence of PEA at 6-HepII in order to serve as an acceptor. Another possibility is that Lpt6 and Lpt3 may be slightly promiscuous at the site of decoration. Although the amino acid sequences

of these two proteins are different, they share common domains that are important for PEA transferase activity. Because the donor and substrate are similar in the PEA modification processes of the 3, 6, and 7 positions of HepII, it is possible that Lpt6 and Lpt3 could act on 3-HepII, 6-HepII, or 7-HepII with differing efficiencies.

Previous studies have shown that different neisserial strains have varying ratios of PEA modification at HepII that do not correspond to the presence or the absence of *lpt3* and *lpt6*. For example, even though *N. meningitidis* strains BZ147 and 2220Y both possess *lpt3* and *lpt6*, BZ147 LOS predominantly contains 3-HepII PEA modification, while 2220Y LOS is simultaneously modified at the 3 and 6 positions of HepII with PEA (22,102). Also, it has been noted that meningococcal strain NMB contains a *lpt3* homolog, but only expresses a LOS with PEA decorated at 3-HepII when a truncated mutant, NMBlgtK, was created and expresses that major LOS form of Hep2KDO2-lipid A. Rather, majority of the LOS displays an L2 immunotype with glucose at 3-HepII and PEA at 6-HepII (69).

Because the presence of *lpt3* does not directly correlate to PEA modification at 3-HepII, as in the example described above, regulation of modification must occur. One process that affects the ability of PEA to add to 3-HepII is the activity of LgtG. It is known that Lpt3 competes with LgtG for the 3-HepII (10). However, this competition between enzymes has complicating factors. First, it has recently been shown that environmental stimulation through the MisS/MisR system alters the level of LgtG activity (130). When the system is not-functional LgtG is upregulated and a higher percentage of LOS molecules are modified with glucose at 3-HepII (69). However, it is important to note that when a NMB *lgtG* mutant was created the resulting LOS still failed

to contain PEA at 3-HepII although the strain possesses *lpt3*. This data suggests that PEA modification at HepII also reflects the functionality of the Lpt3 enzyme. Lpt3 encoded by NMB may have a decreased affinity for either the donor or the substrate in comparison to FA1090 Lpt3. Further evaluation of enzymes from neisserial strains known to have varying ratios of PEA modification of HepII needs to be undertaken to fully understand the modification process of HepII of neisserial LOS.

CHAPTER 7. EVOLUTIONARY ANALYSIS OF NEISSERIAL Lpt3

INTRODUCTION

Neisserial LOS is expressed on the cell surface. Certain structures can provide a selective advantage to the gonococcus in terms of adherence to, invasion of, and survival within host cells. For example, *N. gonorrhoeae* expressing a lacto-*N*-neotetraose structure of LOS preferentially invades human cervical ME180 cells (121). Gonococcal strains expressing terminal *N*-Acetyllactosamine repeats are more infectious than isogenic strains lacking this repeat (64). Sialyation of LOS with CMP-NANA imparts increased serum resistantance for the both the gonococcus and the meningococcus (63). Also, different LOS structures were found on *N. meningitidis* isolates taken from a patient's nasopharyx and the blood (11).

Although, neisserial LOS biosynthetic genes are conserved in the gonococcus and meningococcus, evolution of genes involved in the synthesis of neisserial cell surface structures has been shown to occur. For example, gonococcal isolates taken from patients in the Baltimore area over the last fifteen years revealed that the porin, PorB has been under selection by the immune system. Specific surface-exposed loops of PorB were positively selected and this evolution was also shown in the change of frequency of *porB* alleles (93). Pilin is also another surface antigen found on the *Neisseria*. Certain pilE alleles were found to be under positive selection as particular variants became more frequent in a population (4). Iron aquistion is crucial for neisserial pathogenesis and consequently proteins that bind iron while avoiding an immune response provide the *Neisseria* with a selective advantage. A meningococcal outbreak in West Africa was attributed to the emergence of a mutant *tbpB* allele that encoded a transferrin binding

protein that showed a decreased reaction with the immune system (80). These examples show that upon interaction with the host, certain types of neisserial surface molecules can provide the bacteria with a selective advantage.

Neisseria producing LOS modified with PEA is subject to selective pressure by the immune system. For example, loss of PEA at 3-HepII on *N. meningitidis* LOS resulted in an increased resistance to bactericidal killing and opsonophagocytosis by mAb B5 (81). Also, HepII modified with PEA interacts with C4b of the complement system. It is thought that since PEA at 6-HepII binds C4b more effectively than 3-HepII, strains producing LOS with the 6 modification are more serum sensitive and therefore are less prevalent in the population (102). Because PEA decorated HepII is subject to immune pressures that change the frequency of these modifications in the population, it is possible that variation may exist in the genes that encode the phosphotransferases.

Evidence for variation within the phosphotransferase genes is found in examination of LOS produced by several strains of *Neisseria*. The presence of the *lpt3* gene does not directly correlate with the addition of PEA to 3-HepII of LOS. *N. meningitidis* NMB produces LOS without PEA at 3-HepII even though the strain contains *lpt3* (130). However, when a truncated mutant of this strain was constructed, NMBlgtK (Hep2Kdo2-lipid A), LOS bearing PEA at 3-HepII was observed. Therefore, *lpt3* of this strain may encode a poorly functional Lpt3 transferase (69). Also, *N. meningitidis* strain 89I synthesis LOS that is devoid of PEA at 3-HepII. However, PCR analysis of DNA encoding Lpt3 isolated from this strain resulted in the production of an *lpt3* product that was 8% divergent other *N. meningitidis* strains (preliminary findings of this work). Therefore it was hypothesized that these sequence variations may alter the biochemical

properties of Lpt3 enough to produce the described measurable phenotypic differences. To determine the extent to which this gene is subject to evolutionary forces, sequencing analysis of Lpt3 from a diverse collection of strains was performed.

RESULTS

PCR amplifications of *neisserial lpt3*.

To determine the extent of sequence diversity in *lpt3* an analysis of this gene was performed. Approximately 50 strains, listed in table 1, were chosen for study based on several criteria. First, strains were selected from patients exhibiting different gonococcal disease outcomes such as DGI, PID, and UG. Second, strains were selected based on the site of isolation including the cervix, urethra, pharynx, rectum, synovial fluid, and blood. This collection of *N. gonorrhoeae* reflects the diverse host environments in which the bacteria can survive and the extent to which the gonococcus can cause harm to the host. Third, the collection contains gonococcal isolates collected from STD clinics in a variety of locations to increase the possibility that the isolates are not related (44). Therefore, if LOS with PEA decoration at 3-HepII is exposed to selective forces by the host, this collection would reveal the resulting diversity within Lpt3.

DNA was prepared from each strain in the collection by lysing colonies with sodium hydroxide prior to the addition of neutralizing Tris-HCl. PCR amplification of *lpt3* from these strains were performed using primers PEA-1 and PTE-5. All reactions yielded products consistent with the expected size of *lpt3*, 1500 bp, regardless of the site of isolation or the resulting gonococcal disease.

Nucleotide sequence alignments of *lpt3* in *N. gonorrhoeae.*

Since differences in *lpt3* were not apparent upon agarose gel electrophoresis, the DNA sequence of the amplicons were determined. The resulting sequences are shown in Appendix A. The sequences are grouped according to the disease experienced by the patient from whom the strain was isolated. The *lpt3* nucleotide sequences that were amplified from DGI causing *N. gonorrhoeae* strains were aligned. An initial visual inspection of the alignment reveals two areas within the gene. One portion of the gene, the first 180 nucleotides, contains several sites over which there is a high degree of sequence divergenece. The remainder of the gene is composed of largely conserved sequence. Neisserial housekeeping genes, such as *argJ,* that are not exposed to the host environment characteristically vary intraspecifically between 0 - 1.2 % (34,119). The *lpt3* nucleotide diversity over the first 180 residues is greater than the housekeeping value, and consistent with genes that encode protein products that are selected by evolutionary pressure, such as *porB*, *opa*, and genes enabling antibiotic resistance (30,57,120).

The nucleotide alignment also reveals that two categories of *lpt3* exist within the strains isolated from the DGI patients. One group contains strains DGI 43, DGI Aj, DGI Kb, DGI 3, and FA1090. This group includes a region that could have resulted from homologous recombination, residues 50 through 84, when compared to the corresponding residues from the other DGI strains. The strains not contained within the stated group share a high degree of similarity amongst themselves within the 50 though 84 region.

To determine of the two groups of *lpt3* genes was confined to the DGI strains, nucleotide alignments were also performed with sequences that were derived from strains responsible for PID and UG. As in the DGI strains, significant variation occurs at the 5'

end of *lpt*3. Analysis of the sequence between base pairs 50 and 84 yields two groups similar to those seen in the DGI strains.

Evolution of *lpt3* by phylogenetic analysis.

Phylogenetic analysis of *lpt3* was performed to establish the relationships between *lpt3* from DGI, PID, UG causing *N. gonorrhoeae* strains. The alignments generated from the ClustalW alignments in the previous section were used to create phylogenetic trees using the PHYLIP program. The tree shown in Fig. 43 was generated from the nucleotide alignment of *lpt3* from *N. gonorrhoeae* isolated from patients with DGI, PID, and UG. Analysis of the tree confirms that there are not specific *lpt3* alleles that can differentiate isolates from different gonococcal disease outcomes. For example, strain DIG 43, PID 334, and 517109 SGI were respectively isolated from patients with DGI, PID, and UG disease outcomes. These strains are closely related on the phylogenetic tree. A more detailed inspection of the tree reveals two groupings. The top of the tree contains strains, PID 305, 510 SGI, PID 036, DGI 48, PID 023, DGI 41, 131 SGI, DGI 11 and PID 332. The members of this group contain one form of sequence between nucleotides 50 and 84. These results indicate that sequence divergence at the 5' end of *lpt3* is significant and this site warrants further investigation at the protein level.

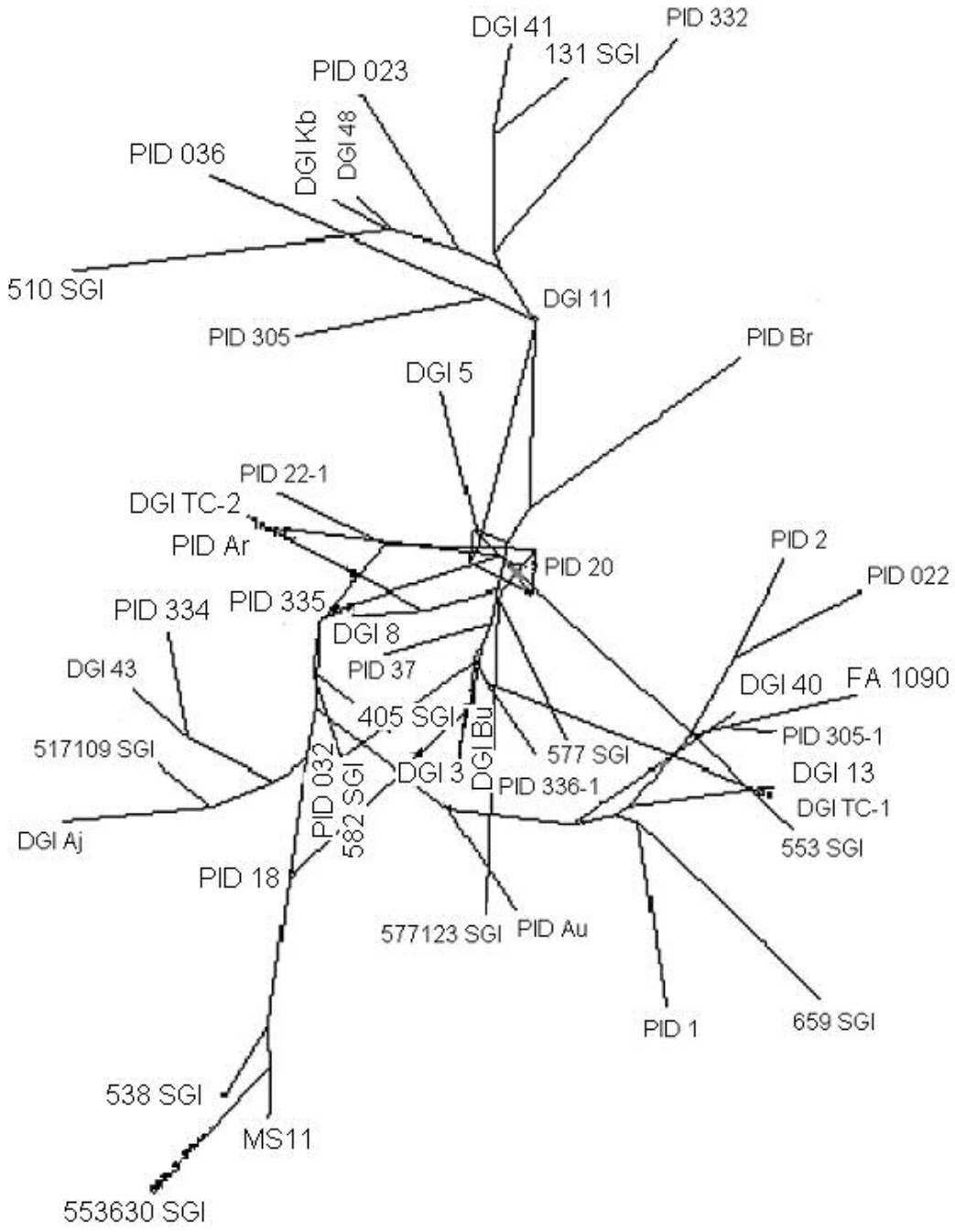


Figure 43. Phylogenetic tree of *lpt3* from *N. gonorrhoeae*.

Amino acid sequence alignments of Lpt3 in pathogenic *Neisseria*.

To determine if the sequence divergence seen at the nucleotide level is translated to the protein level, the *lpt3* sequences were translated using the six-frame translator at the BCM search launcher webite. The resulting Lpt3 amino acid sequences were aligned and grouped according to the disease outcome of the respective strain. These alignments are shown in Appendix B. Lpt3 sequences from strains isolated from patients with DGI are listed first, following Lpt3 sequences from strains isolated from patients with PID and UG.

The results of the Lpt3 alignments show that the diversity seen at the 5' end of *lpt3* is translated to Lpt3. For example, significant residue differences occur between amino acids 12 – 28 regardless of the disease state of the patient from whom the isolate was taken. Other residue changes are scattered throughout the remainder of the protein.

A closer examination of the Lpt3 sequences reveals that there are two different domains in the Lpt3 protein sequences. One Lpt3 domain is demonstrated by isolate DGI 41 and includes amino acids 12 – 28 of SVYTDNDKFKWILRPR. The other type of Lpt3 is exemplified by isolate DGI 43 and contains the sequence PVYYANDISYRFVCGI at those same sites. When these two protein sequences are compared, major residue substituitions are seen. For example, threonine (T) and tyrosine (Y) are respectively contain polar uncharged and aromatic R groups. A change from aspartate (D) to alanine (A) is a change from a negatively charged R group to a nonpolar aliphatic group. Lysine (K) and isoleucine (I) respectively contain R groups that are positively charged and nonpolar. A change from phenylalanine (F) to serine (S) represents a change from an aromatic amino acid to a polar amino acid. Lysine (K) and

tyrosine (Y) respectively contain positively charged and aromatic R groups. A change from threonine (T) to arginine (R) represents a change from polar to a positively charged R group. Isoleucine (I) and phenylalanine (F) are amino acids respectively members of the nonpolar and aromatic groups. A change from arginine (R) and cysteine (C) is a change from a positively charged R group to a polar R group. Prolone (P) and glycine (G) are respectively members of polar and nonpolar R groups. A change from arginine (R) to isoleucine (I) is a change from a positively charged to a nonpolar charged R group. These amino acid differences represent major differences that could affect the Lpt3 function, supporting the idea that two different forms of this protein exist.

Besides the end of Lpt3, there is another site that contains significant diversity seen only in isolates from DGI patients. Lpt3 produced by isolates DGI 13, DGI Bu, DGI 5, DGI 11, DGI 4, DGI 8, and DGI TC-2 contain glutamine (Q) at residue 392 instead of lysine (K). This amino acid substitution results in the loss of a positive charge at that site.

Other nucleotide sequence divergence that was found in the downstream portions of *lpt3* is also reflected at the amino acid level. For example, stop codons are found at several different sites within Lpt3 produced by strains DGI Kb, DGI 43, PID Au, PID 334, 577123 SGI, 582 SGI, PID 023, and 538 SGI. These strains most likely fail to produce a functional Lpt3 transferase.

Evolution of Lpt3 by phylogenetic analysis.

Since the amino acid alignments indicated that there was significant diversity in Lpt3, a phylogenetic analysis was performed. The first 60 amino acids of Lpt3 from gonococcal strains isolated from patients with all three disease outcomes were aligned using the ClustalW algorithm at the Biology Workbench website. From this alignment, a rooted phylogenetic tree was generated and is shown in Fig 44.

Analysis of the tree shows that two distinct evolutionary groups of Lpt3 are present. One group contains Lpt3 from isolates 510 SGI, DGI TC-1, PID 023, PID 336-1, PID 032, DGI Bu, DGI 13, PID 18, PID 37, PID 036, DGI 5, DGI 48, DGI 41, DGI 8, DGI 4, DGI TC-2, 553 SGI, PID 332, PID 20, PID Br, DGI 11, PID 335, PID 305, DGI 40, 577123 SGI, 131 SGI, and 577 SGI. The other group contains Lpt3 from isolates PID 022, PID 2, FA1090, PID 305-1, DGI Kb, PID 1, 405 SGI, DGI 3, 553630 SGI, 538 SGI, PID Au, 517109 SGI, DGI Aj, PID 334, DGI 43, 659 SGI, 582 SGI, and PID 22-1.

Because isolates obtained from all disease types, DGI, PID and UG, express both domains of Lpt3, attention was shifted to the site of isolation of each strain. This step was taken because the neisserial surface is not static. Rather antigenic variation is constantly occurring so that differences may exist between isolates taken from a DGI patient at various sites. Inspection of the isolates comprising each group of Lpt3 reveals a correlation between the isolation site and the Lpt3 expressed by that isolate. Strains DGI 13, DGI 5, DGI 4, and DGI 8 were all isolated from the blood of DGI patients. The phylogenetic tree shows that each of these isolates synthesize Lpt3 of the same type. Also, majority of the isolates taken from the synovial fluid of DGI cases are of this same Lpt3 type. Inspection of other isolation sites including the cervix, urethra, pharynx, and rectum shows that isolates expressing either form of Lpt3 can exist in these tissues.

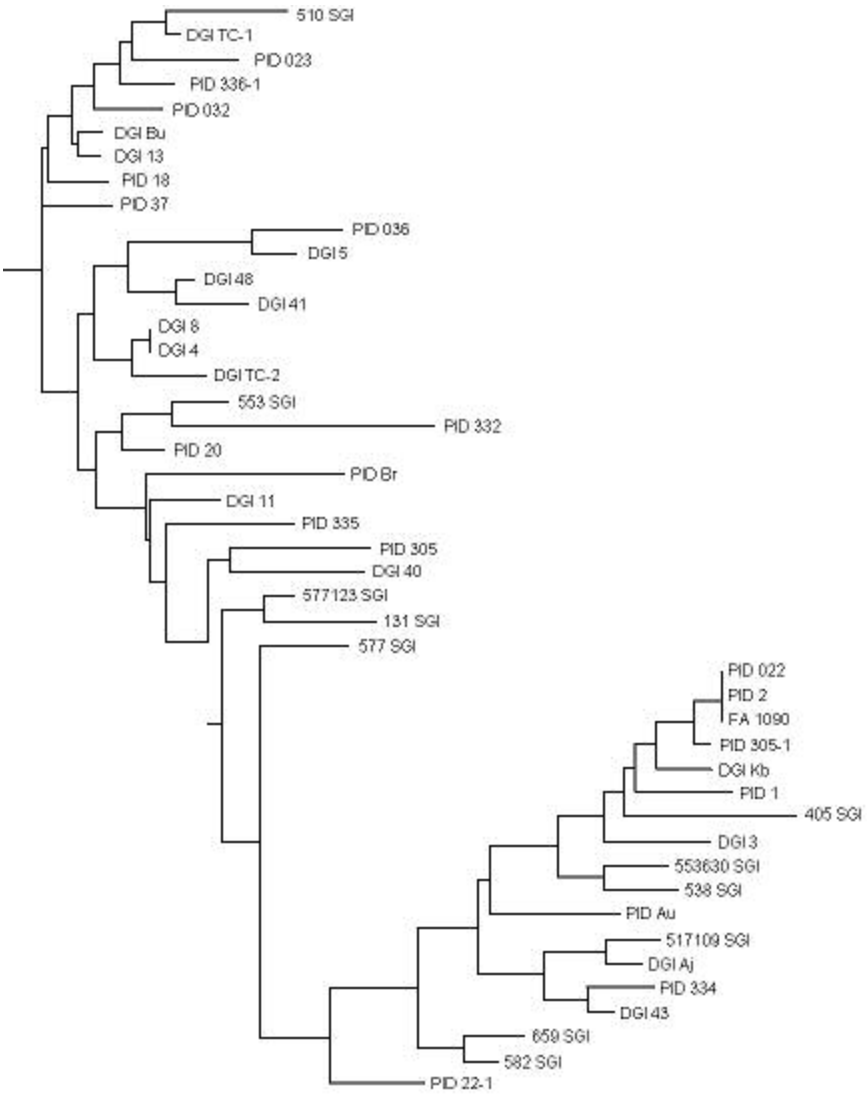


Figure 44. Phylogentic tree of Lpt3 from *N. gonorrhoeae.* The first 60 amino acids of each Lpt3 were aligned using the ClustalW algorithm. Biology Workbench was used to create the rooted tree.

DISCUSSION

In this study, *lpt3* was analyzed across a broad range of pathogenic *Neisseria*. It was necessary to perform this research for several reasons. First, it was reported that Southern hybridization failed to show the presence of *lpt3* in *N. gonorrhoeae* FA1090 although it was identified in this work (81). Therefore, failure of the meningococcal *lpt3* probe to bind FA1090 *lpt3* showed that sequence diversity was likely present in neisserial *lpt3*. Second, it has been reported that the presence of *lpt3* does not correlate with PEA decoration at 3-HepII of LOS (22,68,100). It was determined that sequence analysis should be undertaken to gain insight to the Lpt3 PEA transferase ability.

Nucleotide sequence analysis revealed diversity in *lpt3* that is translated to the protein level. These *lpt3* differences are consistent with the protein being subject to evolutionary forces. It is known that PEA modified LOS is exposed to the host immune response as the complement system more frequently binds to *N. meningitidis* expressing LOS with PEA at 6-HepII than 3-HepII (102). As a result of the increased binding of C4b, lysis of cells producing LOS with PEA at 6-HepII occurs. The sequence diversity that was found in *lpt3* shows that not only is the 6-position exposed to the host immune environment, but that PEA at 3-HepII could also affected.

Phylogenetic analysis of Lpt3 showed the existence of two distinct domains in Lpt3 that differ from each other by a block of non-synonomous amino acid changes between residues 12 and 28. Study of the gonococcal isolates contained in either of the two groups show that the produced form of Lpt3 might influence the disease outcome resulting from gonococcal infection. For example, isolates obtained form the blood and most isolates from the synovial fluid of DGI patients express one type of Lpt3. This

grouping indicates that strains producing Lpt3 have the propensity to disseminate into the blood and the synovial fluid. It is important to note that not all isolates expressing this Lpt3 type caused disseminated infection. It is therefore likely that another factor in conjunction with Lpt3 yields serum resistance to the gonococcus.

The other type of Lpt3 includes isolates obtained from the cervix, rectum, pharynx, and urethra. This group contains isolates that have been described as causing disseminated infection. However, a gonococcal case can be caused by simultaneous infections with multiple strains. Therefore, isolates obtained from the cervix of a DGI patient may not represent the actual disseminated isolate, making the site of isolation more important in these analyses than the type of gonococccal disease experienced by the patient. This observation is significant in that the common laboratory strains, FA1090, F62, MSII and PID2, used in gonococcal research all encode Lpt3 that was not seen in the majority of the blood and synovial fluid isolates. When studying the gonococcal dissemination process, strains expressing the other Lpt3 protein should be selected.

Differences in the enzymatic abilities of the two differing domains of Lpt3 remain unclear. It was previously found that the addition of PEA at 3-HepII to a strain already expressing PEA at 6-HepII confers an increased serum resistance (102). It is likely that a highly active Lpt3 would increase the ratio of LOS molecules containing the decoration at 3-HepII, thereby decreasing the serum susceptibility. Since the isolates from blood and synovial fluid are serum resistant, the form of Lpt3 produced by these cells might be the more active version of the protein. These different activity levels of the two domains could account for the fact that *lpt3* is present in strains such as *N. meningitidis* NMB, while little Lpt3 acitivity is observed (69).

The enzymatic abilities of the Lpt3 proteins require further study. In this work it was determined that PEA transferases may be promiscuous in its site of action in that the enzymes was may act with differing abilities at 3-HepII, 6-HepII, and 7-HepII. It is possible that Lpt3 forms have separate levels of affinity for modification sites on HepII of LOS which further contributes to diversity seen in PEA decoration of HepII. Each type of enzyme should be purified and tested for the ability to modify the different positions of HepII.

CONCLUSIONS

Lipooligosaccharide (LOS) is known to be crucial to neisserial virulence. Specific oligosaccharide structures play a role in tight attachment to epithelial cells (16,121). After invasion, the lipid A portion of the molecule binds TLR4 receptors on macrophages and endothelial cells, resulting in the production of TNF-α and IL1-β, mediators of inflammation (88,125). As a result, inflammation occurs in gonococcal and meningococcal infections. When overproduced in the blood, these factors cause septic shock, intravascular coagulation, and multiple organ failure (5). The OS portion of the molecule also plays a role in the evasion of the host immune system. Both sialyation as well as the location of PEA decoration of the LOS contribute to the prevention of complement-mediated killing (32,85,102). Also, certain oligosaccharide structures can be present in the LOS that mimic antigens found on host cells, protecting the bacteria from an immune response (84).

Because LOS is present on the surface of the *Neisseria* it has been identified as a potential target for neisserial vaccines. In order for a LOS based vaccine to be effective, it needs to be active against all forms of LOS expressed by the *Neisseria*, as this naturally competent genus can exchange DNA encoding different LOS forms. Studies need to be performed to determine these possible structures and the genes required for their synthesis. Until recently, the *Neisseria* field had accepted the statement that this genus expressed LOS, not LPS. As a result, vaccine design proceeded based on these known LOS structures. However, the data presented in this study shows that both genetic and structural diversity is seen in the synthesis of these molecules in the *Neisseria*, therefore additional structures may need to be considered when developing an LOS based vaccine.

On the structural front, this work documents previously undescribed neisserial LOS heterogeneity. MALDI-TOF analysis showed that *N. sicca* 4320 is capable of producing two separate molecules, LOS and LPS. The LOS is predicted to have $(Hex)_2(Hep)_3(PEA)_1(KDO)_1$(lipoidal moiety) as its backbone structure. The accepted composite neisserial LOS structure shown in Fig. 1 contains a diheptose inner core. 4320 also produces a form of LPS that was undefined prior to this work. It consists of an O-antigen [– 3) β-D-GlcNAc (1-X) β-L-Rham (1-] repeat linked to a membrane component of a mass smaller than lipid A. Both 4320 LOS and LPS are novel structures to the *Neisseria*. The structural diversity exhibited by these two molecules is consistent with previous observations. A study of LOS produced by other commensal *Neisseria* showed that while these species produce LOS, these molecules fail to express the necessary epitopes to react with Mabs specific for previously determined LOS structures (6). Also, *N. gonorrhoeae* MS11 mkC is able to produce LOS with additional polylactosamine repeats (64). As a whole, these studies show that the *Neisseria* are capable of synthesizing diverse LOS/LPS molecules, beyond the composite structure that is currently accepted in the field.

This work also shows that significant genetic diversity exists within neisserial LOS biosynthesis genes. The bioinformatic analysis of the *N. sicca* 4320 chromosome revealed regions that show relatively low homology to *N. meningitidis* MC58 *lgtB* and *lgtG*. This low homology to the characterized pathogenic LOS biosynthetic genes is consistent with the literature. Other studies have found nucleotide differences when comparing LOS genes in pathogenic and commensal *Neisseria* (141) (6). These low homology values are expected because the LOS expressed by many commensal strains is

different from the composite pathogenic LOS. As a result, differences are expected in the glycosyl transferases necessary to make the molecules.

Genetic differences are also evident within the neisserial rhamnose biosynthetic cluster. Southern hybridization data shows that the *rfbBAD* cluster in the gonococcus is not identical to the cluster within *N. sicca* 4320, as the later strain showed no hybridization signal for *rfbD*. It is predicted that the rhamnose biosynthetic cluster does not encode functional proteins, since rhamnose is not found in gonococcal LOS (107). The differences found in the 4320 cluster could account for the presence of rhamnose in the O-antigen of the LPS.

Heterogeneity was also found in genes encoding phophoethanolamine transferases, which decorate neisserial LOS with PEA residues. In this work it was biochemically shown that Lpt3 functions as a PEA transferase at 3-HepII of LOS. It has also been speculated that Lpt6 modifies 6-HepII with PEA (139). The genes encoding both of these proteins were also shown to contain nucleotide diversity. Southern hybridization analysis of *lpt3* and *lpt6* across the genus shows that although these genes are present in the gonococcal and meningococcal strains, they have a variable presence in the *N. sicca* strains analyzed. Even within the gonococcus, diversity was observed within Lpt3. The diversity is apparent even at the protein level as two forms of enzyme exist.

The heterogeneity of the LOS/LPS genes in the *Neisseria* may make a difference in the genus' pathogenic ability. For example, *N. sicca* 4320 was the causative agent of a fatal endocarditis case in an otherwise healthy adult male. It is possible that both the LOS and LPS structure contribute to its increased virulence. In regards to LOS, it is known that the presence of PEA at 6-HepII of LOS allows for the molecule to serve as a receptor

for complement component C4b. Therefore, strains with PEA at 3-HepII have an increased resistance to complement killing (102). The presence of *lpt3* in *N. sicca* 4320, as indicated by Southern hybridization, might contribute to the serum resistance of the 4320 strain. The presence of LPS, increases the length of this type of molecule which may hinder attachment of complement components, much like the meningococcal capsule (51).

Another example of how neisserial virulence can be altered by the expressed LOS is demonstrated by the evolutionary analysis of gonococcal Lpt3. Two differing domains of Lpt3 were evident upon phylogenetic analysis of isolates from different sites of the host. Isolates that had disseminated into the bloodstream contain the genetic information to express one of the two Lpt3 domains. The difference in activity and specificity of the two Lpt3 domains might have an influence on the disease outcome of a gonococcal infection.

The results of this study show that the specific structure of neisserial LOS can be enough to alter the pathogenicity of the *Neisseria*. *N. gonorrhoeae* is an obligate human pathogen. The data presented in this study shows that *lpt3* is present in all of the strains examined. *N. meningitidis* is an opportunisitic pathogen. Other studies have shown that *lpt3* is not present in all meningococcal strains (81). The data contained in this study show that *lpt3* is not present in strains of *N. sicca* that have not caused disease. However, both strains that have been isolated from fatal cases of endocarditis, 4319 and 4320, contain *lpt3*. Therefore, the data suggests that PEA decoration at 3-HepII of neisserial LOS is able to make a commensal *Neisseria* pathogenic.

This work shows that diversity exists in the neisserial LOS/LPS beyond what is currently accepted in the literature. Additonally, this study raises several avenues for future research. The most important areas to pursue involve the identification of the LPS synthesis genes of *N. sicca* 4320 and testing the activity of the Lpt3 enzymes. The rhamnose synthesis genes of 4320 should be inactivated to determine if the resulting LPS is altered. Also, the 4320 potential *lgtB* and *lgtG* homologs should be mutated to observe an affect on either the LOS or LPS molecules. It would also be interesting to construct an *lpt3* mutant of 4320 to show that the LOS is decorated with PEA at 3-HepII. Futher structural work is required on the 4320 LPS. MALDI-MS analysis should be performed with a lower mass gate on the purified LPS molecule to determine the size of the O-repeat anchor. Also, based on the Lpt3 protein divergence seen in the gonococcus, the different domains of Lpt3 need to be examined for their contributions to the enzyme activity.

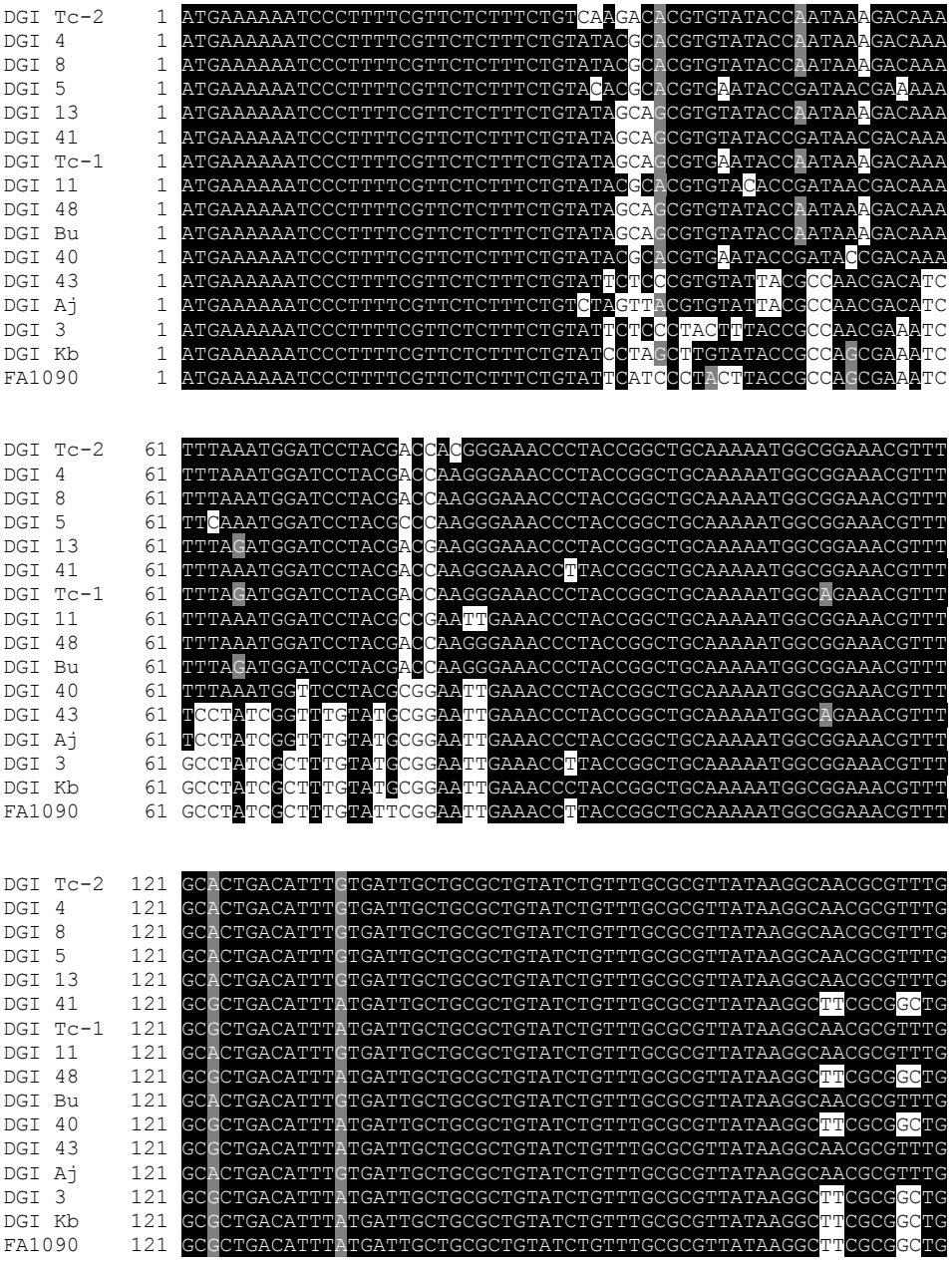
If this study were to be undertaken again, several changes in methodology would be appropriate. For example, LOS and LPS mutants made in 4320 should not be constructed using a kanamycin resistance marker. The absence of LPS from the 4320 membrane likely affects the organism's resistance to kanamycin. In addition, for strucutural studies, 4320 LOS and LPS should be separated by size exclusion chromatography to prior to analysis of either molecule. Separation of the molecule would allow for clearer mass spec and gas chromatography analysis.

The results of this study allow for the development of specific aims for future research. One major goal is to determine if the 4320 LPS synthesis genes are present in other *Neisseria*. The presence of these genes in the pathogenic *Neisseria* strains would

indicate that LPS could be produced in-vivo, increasing neisserial virulence. Another aim of future research is to determine the affinities of the Lpt3 and Lpt6 transferase genes for PEA decoration at 3, 6, and 7 positions of 3-HepII. Demonstrating promiscuity of these enzymes will help to explain the differing ratios of PEA decoration of neisserial LOS. Finally, PEA modification of 3-HepII of LOS needs to be examined for its role in altering neisserial pathogenicity. Experiments could be performed to compare the pathgencity of *N. sicca* 4320 to an isogenic strain lacking the *lpt3* gene.

Appendix A: Nucleotide sequences of *lpt3* from *N. gonorrhoeae*.

Shown are the *lpt3* nucleotide alignments from isolates taken from patients with DGI. Shaded regions indicate areas of conservation. Gray and white boxes indicate where nucleotide differences exist.



```
DGI Tc-2  181 TTGATTGCGGTGTTTTTCGCGTTCAGCATTATTGCCAACAATGTGCATTACGCGGTTTAT
DGI 4     181 TTGATTGCGGTGTTTTTCGCGTTCAGCATTATTGCCAACAATGTGCATTACGCGGTTTAT
DGI 8     181 TTGATTGCGGTGTTTTTCGCGTTCAGCATTATTGCCAACAATGTGCATTACGCGGTTTAT
DGI 5     181 TTGATTGCGGTGTTTTTCGCGTTCAGCATTATTGCCAACAATGTGCATTACGCGGTTTAT
DGI 13    181 TTGATTGCGGTGTTTTTCGCGTTCAGCATTATTGCCAACAATGTGCATTACGCGGTTTAT
DGI 41    181 CTGATTGCGGTGTTTTTCGCGTTCAGCATGATTGCCAACAATGTGCATTACGCGGTTTAT
DGI Tc-1  181 TTGATTGCGGTGTTTTTCGCGTTCAGCATTATTGCCAACAATGTGCATTACGCGGTTTAT
DGI 11    181 TTGATTGCGGTGTTTTTCGCGTTCAGCATTATTGCCAACAATGTGCATTACGCGGTTTAT
DGI 48    181 CTGATTGCGGTGTTTTTCGCGTTCAGCATGATTGCCAACAATGTGCATTACGCGGTTTAT
DGI Bu    181 TTGATTGCGGTGTTTTTCGCGTTCAGCATTATTGCCAACAATGTGCATTACGCGGTTTAT
DGI 40    181 CTGATTGCGGTGTTTTTCGCGTTCAGCATGATTGCCAACAATGTGCATTACGCGGTTTAT
DGI 43    181 TTGATTGCGGTGTTTTTCGCGTTCAGCATTATTGCCAACAATGTGCATTACGCGGTTTAT
DGI Aj    181 TTGATTGCGGTGTTTTTCGCGTTCAGCATTATTGCCAACAATGTGCATTACGCGGTTTAT
DGI 3     181 CTGATTGCGGTGTTTTTCGCGTTCAGCATGATTGCCAACAATGTGCATTACGCGGTTTAT
DGI Kb    181 CTGATTGCGGTGTTTTTCGCGTTCAGCATGATTGCCAACAATGTGCATTACGCGGTTTAT
FA1090    181 CTGATTGCGGTGTTTTTCGCGTTCAGCATGATTGCCAACAATGTGCATTACGCGGTTTAT

DGI Tc-2  241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
DGI 4     241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
DGI 8     241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
DGI 5     241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
DGI 13    241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
DGI 41    241 CAAAGCTGGATGACGGGCATCAATTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
DGI Tc-1  241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
DGI 11    241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCAAAGTCGGC
DGI 48    241 CAAAGCTGGATGACGGGCATCAATTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
DGI Bu    241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
DGI 40    241 CAAAGCTGGATGACGGGCATCAATTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
DGI 43    241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
DGI Aj    241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
DGI 3     241 CAAAGCTGGATGACGGGCATCAATTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
DGI Kb    241 CAAAGCTGGATGACGGGCATCAATTATTAGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
FA1090    241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC

DGITc-2   301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
DGI 4     301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
DGI 8     301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
DGI 5     301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
DGI 13    301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
DGI 41    301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
DGITc-1   301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
DGI 11    301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
DGI 48    301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
DGI Bu    301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
DGI 40    301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
DGI 43    301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
DGI Aj    301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
DGI 3     301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
DGI Kb    301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
FA1090    301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA

DGI Tc-2  361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
DGI 4     361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
DGI 8     361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
DGI 5     361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
DGI 13    361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
DGI 41    361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
DGI Tc-1  361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
DGI 11    361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
DGI 48    361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
DGI Bu    361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
DGI 40    361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
DGI 43    361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
DGI Aj    361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
DGI 3     361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
DGI Kb    361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
FA1090    361 GTCATGTTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATA
```

```
DGI Tc-2  421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
DGI 4     421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
DGI 8     421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
DGI 5     421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
DGI 13    421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
DGI 41    421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
DGI Tc-1  421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
DGI 11    421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
DGI 48    421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
DGI Bu    421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
DGI 40    421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
DGI 43    421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
DGI Aj    421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
DGI 3     421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
DGI Kb    421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
FA1090    421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC

DGI Tc-2  481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
DGI 4     481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
DGI 8     481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
DGI 5     481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
DGI 13    481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
DGI 41    481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
DGI Tc-1  481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
DGI 11    481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
DGI 48    481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
DGI Bu    481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
DGI 40    481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
DGI 43    481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
DGI Aj    481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
DGI 3     481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
DGI Kb    481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
FA1090    481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT

DGI Tc-2  541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
DGI 4     541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
DGI 8     541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
DGI 5     541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
DGI 13    541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
DGI 41    541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
DGI Tc-1  541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
DGI 11    541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
DGI 48    541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
DGI Bu    541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
DGI 40    541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
DGI 43    541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
DGI Aj    541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
DGI 3     541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
DGI Kb    541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
FA1090    541 TTTGTCGGGCGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA

DGI Tc-2  601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
DGI 4     601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
DGI 8     601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
DGI 5     601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
DGI 13    601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
DGI 41    601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
DGI Tc-1  601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
DGI 11    601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
DGI 48    601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
DGI Bu    601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
DGI 40    601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
DGI 43    601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
DGI Aj    601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
DGI 3     601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
DGI Kb    601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
FA1090    601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
```


DGI Tc-2 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
DGI 4 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
DGI 8 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
DGI 5 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
DGI 13 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
DGI 41 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
DGI Tc-1 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
DGI 11 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
DGI 48 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
DGI Bu 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
DGI 40 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
DGI 43 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
DGI Aj 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
DGI 3 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
DGI Kb 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
FA1090 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT


DGI Tc-2 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
DGI 4 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
DGI 8 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
DGI 5 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
DGI 13 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
DGI 41 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
DGI Tc-1 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
DGI 11 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
DGI 48 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
DGI Bu 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
DGI 40 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
DGI 43 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
DGI Aj 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
DGI 3 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
DGI Kb 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
FA1090 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC


DGI Tc-2 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
DGI 4 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
DGI 8 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
DGI 5 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
DGI 13 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTC
DGI 41 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
DGI Tc-1 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
DGI 11 781 TTTATGACGGCAGTATCCCTTTCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
DGI 48 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
DGI Bu 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
DGI 40 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
DGI 43 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
DGI Aj 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
DGI 3 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
DGI Kb 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
FA1090 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG

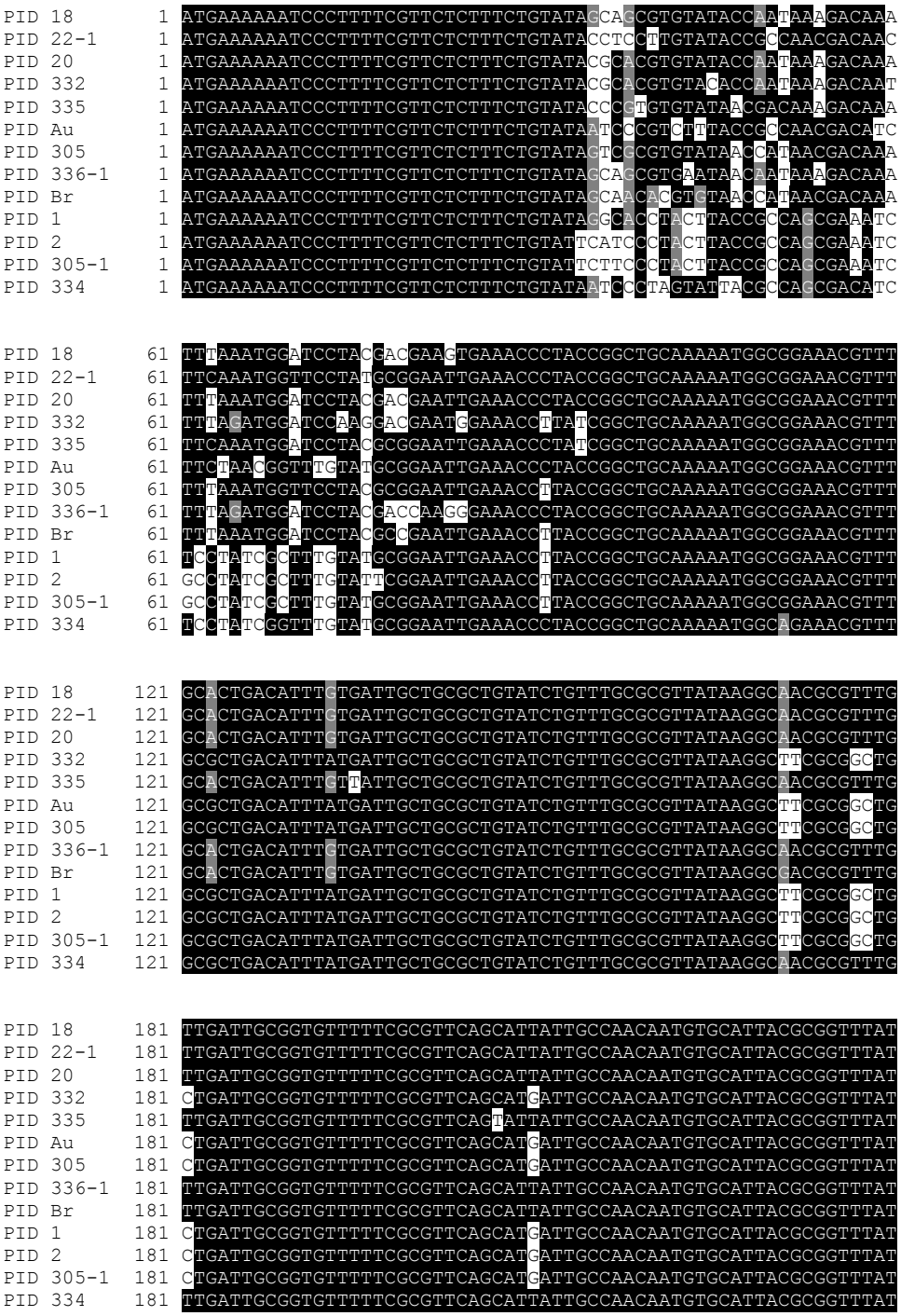

DGI Tc-2 841 TATCAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
DGI 4 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
DGI 8 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
DGI 5 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
DGI 13 841 GTACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
DGI 41 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
DGI Tc-1 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
DGI 11 841 GTACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
DGI 48 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
DGI Bu 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
DGI 40 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
DGI 43 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
DGI Aj 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
DGI 3 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
DGI Kb 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
FA1090 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA

```
DGI Tc-2  901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
DGI 4     901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
DGI 8     901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
DGI 5     901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
DGI 13    901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
DGI 41    901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
DGI Tc-1  901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
DGI 11    901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
DGI 48    901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
DGI Bu    901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
DGI 40    901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
DGI 43    901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
DGI Aj    901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
DGI 3     901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
DGI Kb    901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
FA1090    901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG

DGI Tc-2  961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
DGI 4     961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
DGI 8     961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
DGI 5     961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
DGI 13    961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
DGI 41    961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
DGI Tc-1  961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
DGI 11    961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
DGI 48    961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
DGI Bu    961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
DGI 40    961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
DGI 43    961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
DGI Aj    961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
DGI 3     961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
DGI Kb    961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
FA1090    961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG

DGI Tc-2 1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
DGI 4    1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
DGI 8    1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
DGI 5    1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
DGI 13   1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
DGI 41   1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
DGI Tc-1 1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
DGI 11   1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
DGI 48   1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
DGI Bu   1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
DGI 40   1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
DGI 43   1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
DGI Aj   1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
DGI 3    1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
DGI Kb   1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
FA1090   1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT

DGI Tc-2 1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
DGI 4    1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
DGI 8    1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
DGI 5    1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
DGI 13   1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
DGI 41   1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
DGI Tc-1 1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
DGI 11   1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
DGI 48   1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
DGI Bu   1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
DGI 40   1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
DGI 43   1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
DGI Aj   1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
DGI 3    1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
DGI Kb   1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
FA1090   1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
```

```
DGI Tc-2 1141 AAAGTATTCGGCGAAGCCGATATTGTGGATCAGTACGACAACACCATCCACAAAACCGAC
DGI 4    1141 AAAGTATTCGGCGAAGCCGATATTGTGGATCAGTACGACAACACCATCCACAAAACCGAC
DGI 8    1141 AAAGTATTCGGCGAAGCCGATATTGTGGATCAGTACGACAACACCATCCACAAAACCGAC
DGI 5    1141 AAAGTATTCGGCGAAGCCGATATTGTGGATCAGTACGACAACACCATCCACAAAACCGAC
DGI 13   1141 AAAGTATTCGGCGAAGCCGATATTGTGGATCAGTACGACAACACCATCCACAAAACCGAC
DGI 41   1141 AAAGTATTCGGCGAAGCCGATATTGCGGATAAGTACGACAACACCATCCACAAAACCGAC
DGI Tc-1 1141 AAAGTATTCGGCGAAGCCGATATTGTGGATAAGTACGACAACACCATCCACAAAACCGAC
DGI 11   1141 AAAGTATTCGGCGAAGCCGATATTGTGGATCAGTACGACAACACCATCCACAAAACCGAC
DGI 48   1141 AAAGTATTCGGCGAAGCCGATATTGCGGATAAGTACGACAACACCATCCACAAAACCGAC
DGI Bu   1141 AAAGTATTCGGCGAAGCCGATATTGTGGATCAGTACGACAACACCATCCACAAAACCGAC
DGI 40   1141 AAAGTATTCGGCGAAGCCGATATTGCGGATAAGTACGACAACACCATCCACAAAACCGAC
DGI 43   1141 AAAGTATTCGGCGAAGCCGATATTGTGGATAAGTACGACAACACCATCCACAAAACCGAC
DGI Aj   1141 AAAGTATTCGGCGAAGCCGATATTGTGGATAAGTACGACAACACCATCCACAAAACCGAC
DGI 3    1141 AAAGTATTCGGCGAAGCCGATATTGCGGATAAGTACGACAACACCATCCACAAAACCGAC
DGI Kb   1141 AAAGTATTCGGCGAAGCCGATATTGCGGATAAGTACGACAACACCATCCACAAAACCGAC
FA1090   1141 AAAGTATTCGGCGAAGCCGATATTGTGGATAAGTACGACAACACCATCCACAAAACCGAC

DGI Tc-2 1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
DGI 4    1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
DGI 8    1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
DGI 5    1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
DGI 13   1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
DGI 41   1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
DGI Tc-1 1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
DGI 11   1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
DGI 48   1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
DGI Bu   1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
DGI 40   1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
DGI 43   1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTAGCTGTTT
DGI Aj   1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
DGI 3    1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
DGI Kb   1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
FA1090   1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT

DGI Tc-2 1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
DGI 4    1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
DGI 8    1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
DGI 5    1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
DGI 13   1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
DGI 41   1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
DGI Tc-1 1261 GCCTATACCTCCGATCATGACCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
DGI 11   1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
DGI 48   1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
DGI Bu   1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
DGI 40   1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
DGI 43   1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
DGI Aj   1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
DGI 3    1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
DGI Kb   1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
FA1090   1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG

DGI Tc-2 1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
DGI 4    1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
DGI 8    1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
DGI 5    1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
DGI 13   1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
DGI 41   1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
DGI Tc-1 1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
DGI 11   1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
DGI 48   1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
DGI Bu   1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
DGI 40   1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
DGI 43   1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
DGI Aj   1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
DGI 3    1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
DGI Kb   1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
FA1090   1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
```

```
DGI Tc-2 1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
DGI 4    1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
DGI 8    1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
DGI 5    1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
DGI 13   1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
DGI 41   1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
DGI Tc-1 1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
DGI 11   1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
DGI 48   1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
DGI Bu   1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
DGI 40   1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
DGI 43   1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
DGI Aj   1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
DGI 3    1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
DGI Kb   1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
FA1090   1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC

DGI Tc-2 1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
DGI 4    1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
DGI 8    1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
DGI 5    1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
DGI 13   1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
DGI 41   1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
DGI Tc-1 1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
DGI 11   1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
DGI 48   1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
DGI Bu   1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
DGI 40   1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
DGI 43   1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
DGI Aj   1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
DGI 3    1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
DGI Kb   1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
FA1090   1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA

DGI Tc-2 1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
DGI 4    1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
DGI 8    1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
DGI 5    1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
DGI 13   1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
DGI 41   1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
DGI Tc-1 1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
DGI 11   1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
DGI 48   1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
DGI Bu   1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
DGI 40   1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
DGI 43   1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
DGI Aj   1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
DGI 3    1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
DGI Kb   1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
FA1090   1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT

DGI Tc-2 1561 GTTTATCCGCAATAAGC
DGI 4    1561 GTTTATCCGCATTAAGC
DGI 8    1561 GTTTATCCGCATTAAGC
DGI 5    1561 GTTTATCCGCAATAAGC
DGI 13   1561 GTTTATCCGCAATAAGC
DGI 41   1561 GTTTATCCGCAATAAGC
DGI Tc-1 1561 GTTTATCCGCATTAAGC
DGI 11   1561 GTTTATCCGCAATAAGC
DGI 48   1561 GTTTATCCGCATTAAGC
DGI Bu   1561 GTTTATCCGCAATAAGC
DGI 40   1561 GTTTATCCGCAATAAGC
DGI 43   1561 GTTTATCCGCAATAAGC
DGI Aj   1561 GTTTATCCGCATTAAGC
DGI 3    1561 GTTTATCCGCAATAAGC
DGI Kb   1561 GTTTATCCGAAATAAGC
FA1090   1561 GTTTATCCGCAATAAGC
```

Shown are the *lpt3* nucleotide alignments from isolates taken from patients with PID. Shaded regions indicate areas of conservation. Gray and white boxes indicate where nucleotide differences exist.



```
PID 18     241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
PID 22-1   241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
PID 20     241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
PID 332    241 CAAAGCTGGATGACGGGCATCAATTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
PID 335    241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
PID Au     241 CAAAGCTGGATGACGGGCATCAATTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
PID 305    241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
PID 336-1  241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
PID Br     241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
PID 1      241 CAAAGCTGGATGACGGGCATCAATTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
PID 2      241 CAAAGCTGGATGACGGGCATCAATTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
PID 305-1  241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
PID 334    241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC

PID 18     301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
PID 22-1   301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
PID 20     301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
PID 332    301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
PID 335    301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
PID Au     301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
PID 305    301 AGCGCGGGCGCGTCGATGTTGGATAAATTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
PID 336-1  301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
PID Br     301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
PID 1      301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
PID 2      301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
PID 305-1  301 AGCGCGGGCGCGTCGATGTTGGATAAATTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
PID 334    301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA

PID 18     361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
PID 22-1   361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
PID 20     361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
PID 332    361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
PID 335    361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
PID Au     361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
PID 305    361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
PID 336-1  361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
PID Br     361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
PID 1      361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
PID 2      361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
PID 305-1  361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
PID 334    361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC

PID 18     421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
PID 22-1   421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
PID 20     421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
PID 332    421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
PID 335    421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
PID Au     421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
PID 305    421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
PID 336-1  421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
PID Br     421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
PID 1      421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
PID 2      421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
PID 305-1  421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
PID 334    421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
```

```
PID 18      481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
PID 22-1    481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
PID 20      481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
PID 332     481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
PID 335     481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
PID Au      481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
PID 305     481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
PID 336-1   481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
PID Br      481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
PID 1       481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
PID 2       481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
PID 305-1   481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
PID 334     481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT

PID 18      541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
PID 22-1    541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
PID 20      541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
PID 332     541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
PID 335     541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
PID Au      541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
PID 305     541 TTTGTCGGGCGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
PID 336-1   541 TTTGTCGGGCGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
PID Br      541 TTTGTCGGGCGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
PID 1       541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
PID 2       541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
PID 305-1   541 TTTGTCGGGCGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
PID 334     541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA

PID 18      601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
PID 22-1    601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
PID 20      601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
PID 332     601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
PID 335     601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
PID Au      601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
PID 305     601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
PID 336-1   601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
PID Br      601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
PID 1       601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
PID 2       601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
PID 305-1   601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
PID 334     601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC

PID 18      661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
PID 22-1    661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
PID 20      661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
PID 332     661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
PID 335     661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
PID Au      661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
PID 305     661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
PID 336-1   661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
PID Br      661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
PID 1       661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
PID 2       661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
PID 305-1   661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
PID 334     661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
```

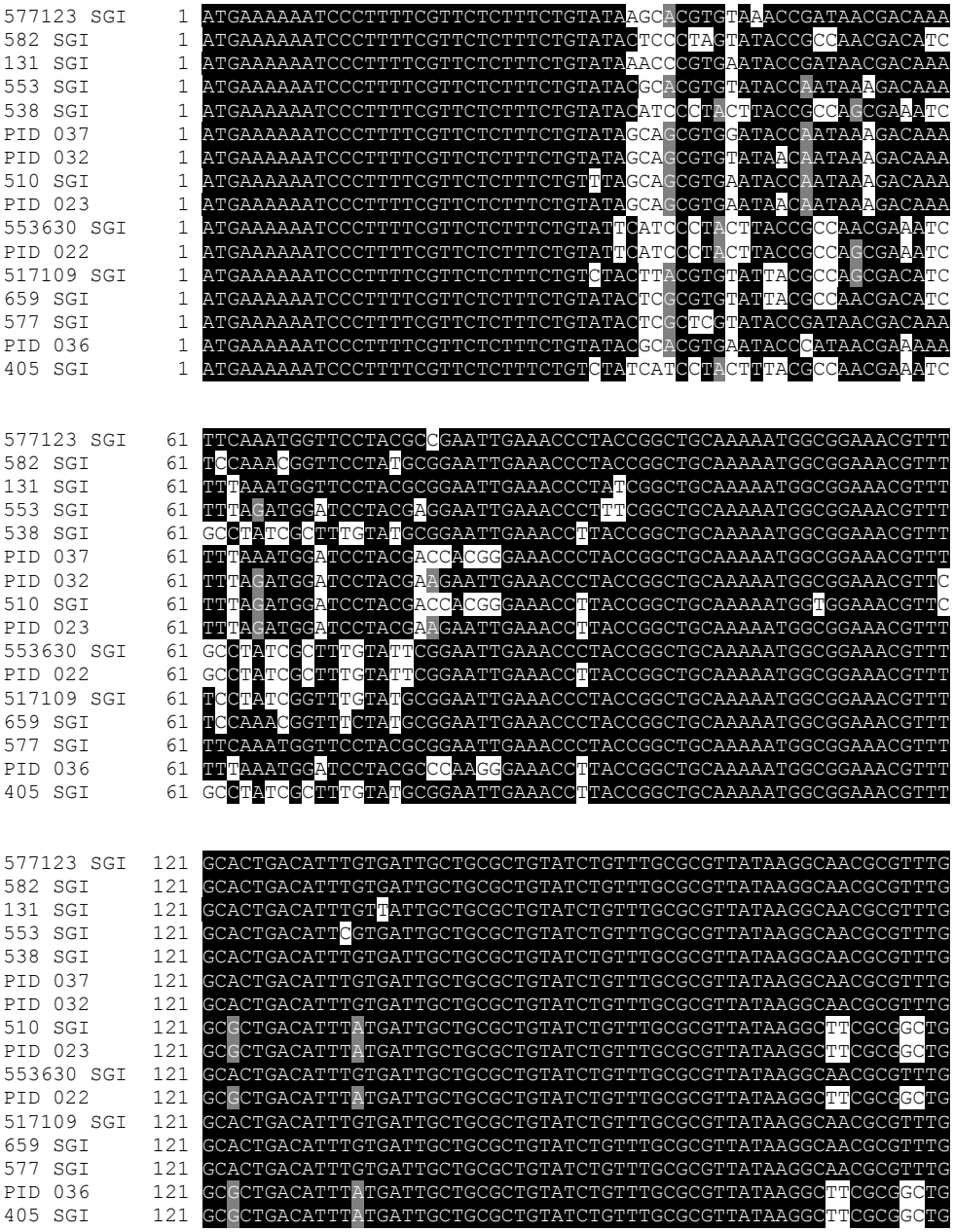
```
PID 18     721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
PID 22-1   721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
PID 20     721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
PID 332    721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
PID 335    721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
PID Au     721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
PID 305    721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
PID 336-1  721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
PID Br     721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
PID 1      721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
PID 2      721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGGGAAACAAAGTTATCCCGCAGGC
PID 305-1  721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
PID 334    721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC

PID 18     781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTC
PID 22-1   781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
PID 20     781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
PID 332    781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
PID 335    781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
PID Au     781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
PID 305    781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
PID 336-1  781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
PID Br     781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
PID 1      781 TTTATGACGGCAGTATCCTTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
PID 2      781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
PID 305-1  781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
PID 334    781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG

PID 18     841 TTACAAAGCAGCGGCGGCGAAACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
PID 22-1   841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
PID 20     841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
PID 332    841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
PID 335    841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
PID Au     841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
PID 305    841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
PID 336-1  841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
PID Br     841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
PID 1      841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
PID 2      841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
PID 305-1  841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
PID 334    841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA

PID 18     901 ACGTATTTTTACACTTTCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
PID 22-1   901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
PID 20     901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
PID 332    901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
PID 335    901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
PID Au     901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
PID 305    901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
PID 336-1  901 ACGTATTTTTACAGTGCTCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
PID Br     901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
PID 1      901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
PID 2      901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
PID 305-1  901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
PID 334    901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
```

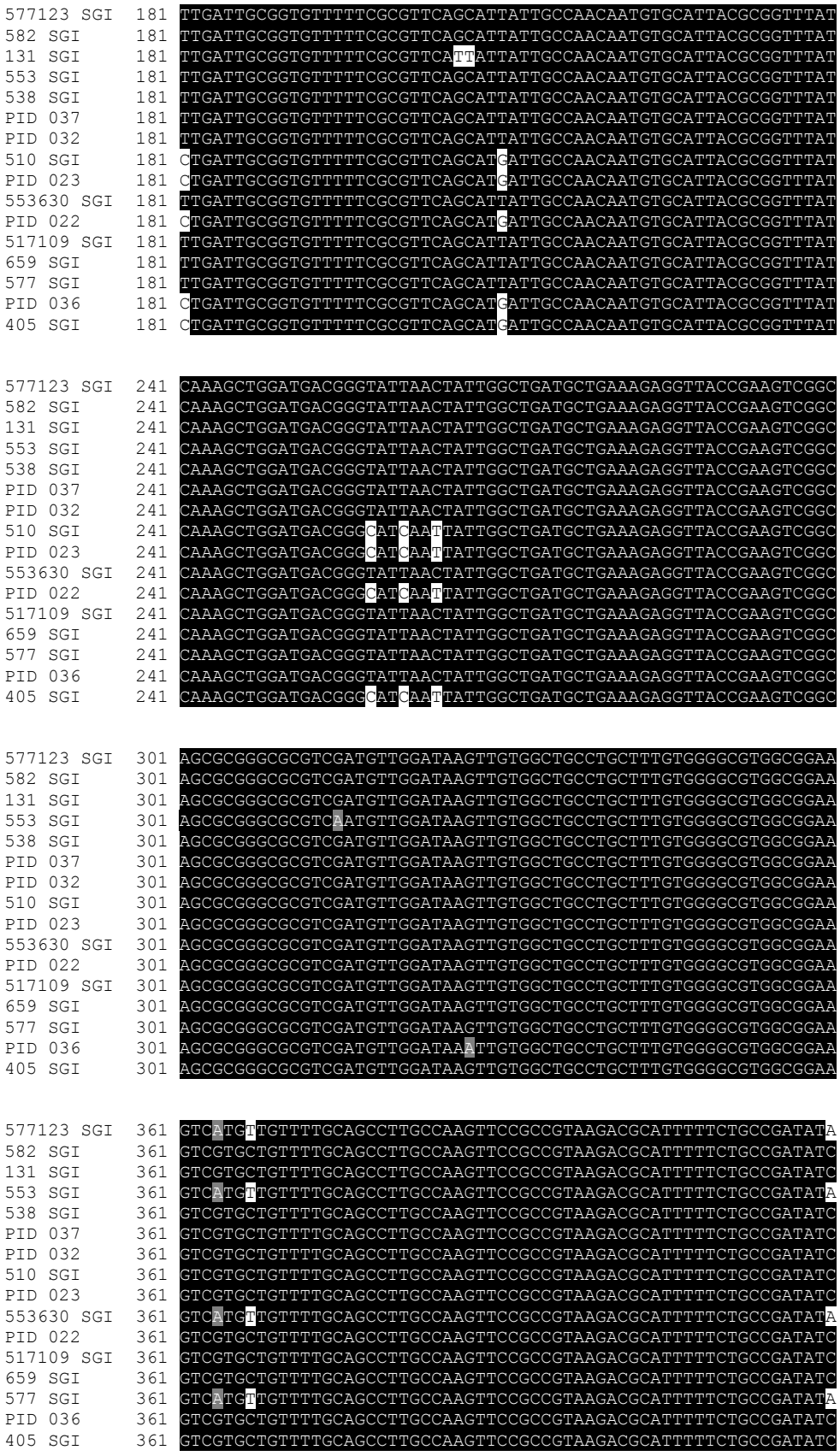
```
PID 18     961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
PID 22-1   961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
PID 20     961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
PID 332    961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
PID 335    961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
PID Au     961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
PID 305    961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
PID 336-1  961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
PID Br     961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
PID 1      961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
PID 2      961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
PID 305-1  961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
PID 334    961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG

PID 18    1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
PID 22-1  1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
PID 20    1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
PID 332   1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
PID 335   1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
PID Au    1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
PID 305   1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
PID 336-1 1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
PID Br    1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
PID 1     1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
PID 2     1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
PID 305-1 1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
PID 334   1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT

PID 18    1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
PID 22-1  1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
PID 20    1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
PID 332   1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
PID 335   1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
PID Au    1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
PID 305   1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
PID 336-1 1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
PID Br    1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
PID 1     1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
PID 2     1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
PID 305-1 1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
PID 334   1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT

PID 18    1141 AAAGTATTCGGCGAAGCCGATATTGTGGATAAGTACGACAACACCATCCACAAAACCGAC
PID 22-1  1141 AAAGTATTCGGCGAAGCCGATATTGTGGATAAGTACGACAACACCATCCACAAAACCGAC
PID 20    1141 AAAGTATTCGGCGAAGCCGATATTGTGGATAAGTACGACAACACCATCCACAAAACCGAC
PID 332   1141 AAAGTATTCGGCGAAGCCGATATTGTGGATAAGTACGACAACACCATCCACAAAACCGAC
PID 335   1141 AAAGTATTCGGCGAAGCCGATATTGTGGATAAGTACGACAACACCATCCACAAAACCGAC
PID Au    1141 AAAGTATTCGGCGAAGCCGATATTGCGGATAAGTACGACAACACCATCCACAAAACCGAC
PID 305   1141 AAAGTATTCGGCGAAGCCGATATTGTGGATAAGTACGACAACACCATCCACAAAACCGAC
PID 336-1 1141 AAAGTATTCGGCGAAGCCGATATTGTGGATAAGTACGACAACACCATCCACAAAACCGAC
PID Br    1141 AAAGTATTCGGCGAAGCCGATATTGTGGATAAGTACGACAACACCATCCACAAAACCGAC
PID 1     1141 AAAGTATTCGGCGAAGCCGATATTGCGGATAAGTACGACAACACCATCCACAAAACCGAC
PID 2     1141 AAAGTATTCGGCGAAGCCGATATTGCGGATAAGTACGACAACACCATCCACAAAACCGAC
PID 305-1 1141 AAAGTATTCGGCGAAGCCGATATTGTGGATAAGTACGACAACACCATCCACAAAACCGAC
PID 334   1141 AAAGTATTCGGCGAAGCCGATATTGTGGATAAGTACGACAACACCATCCACAAAACCGAC
```

```
PID 18    1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
PID 22-1  1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
PID 20    1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
PID 332   1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
PID 335   1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
PID Au    1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
PID 305   1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
PID 336-1 1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
PID Br    1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
PID 1     1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
PID 2     1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
PID 305-1 1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
PID 334   1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT

PID 18    1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
PID 22-1  1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
PID 20    1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
PID 332   1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
PID 335   1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
PID Au    1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
PID 305   1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
PID 336-1 1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
PID Br    1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
PID 1     1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
PID 2     1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
PID 305-1 1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
PID 334   1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG

PID 18    1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
PID 22-1  1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
PID 20    1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
PID 332   1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
PID 335   1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
PID Au    1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
PID 305   1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
PID 336-1 1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
PID Br    1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
PID 1     1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
PID 2     1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
PID 305-1 1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
PID 334   1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG

PID 18    1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
PID 22-1  1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
PID 20    1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
PID 332   1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
PID 335   1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
PID Au    1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
PID 305   1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
PID 336-1 1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
PID Br    1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
PID 1     1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
PID 2     1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
PID 305-1 1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
PID 334   1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
```

```
PID 18    1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
PID 22-1  1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAAGCTCGGTAACA
PID 20    1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
PID 332   1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
PID 335   1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
PID Au    1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
PID 305   1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
PID 336-1 1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
PID Br    1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
PID 1     1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
PID 2     1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
PID 305-1 1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
PID 334   1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA

PID 18    1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
PID 22-1  1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
PID 20    1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
PID 332   1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
PID 335   1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
PID Au    1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
PID 305   1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
PID 336-1 1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
PID Br    1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
PID 1     1501 GGCAACCTGATTACGGGCGATGCAAGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
PID 2     1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
PID 305-1 1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
PID 334   1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT

PID 18    1561 GTTTATCCGCAATAAGC
PID 22-1  1561 GTTTATCCGCAATAAGC
PID 20    1561 GTTTATCCGCAATAAGC
PID 332   1561 GTTTATCCGCAATAAGC
PID 335   1561 GTTTATCCGCAATAAGC
PID Au    1561 GTTTATCCGCAATAAGC
PID 305   1561 GTTTATCCGCAATAAGC
PID 336-1 1561 GTTTATCCGCAATAAGC
PID Br    1561 GTTTATCCGCAATAAGC
PID 1     1561 GTTTATCCGCATTAAGC
PID 2     1561 GTTTATCCGCATTAAGC
PID 305-1 1561 GTTTATCCGCAATAAGC
PID 334   1561 GTTTATCCGCAATAAGC
```

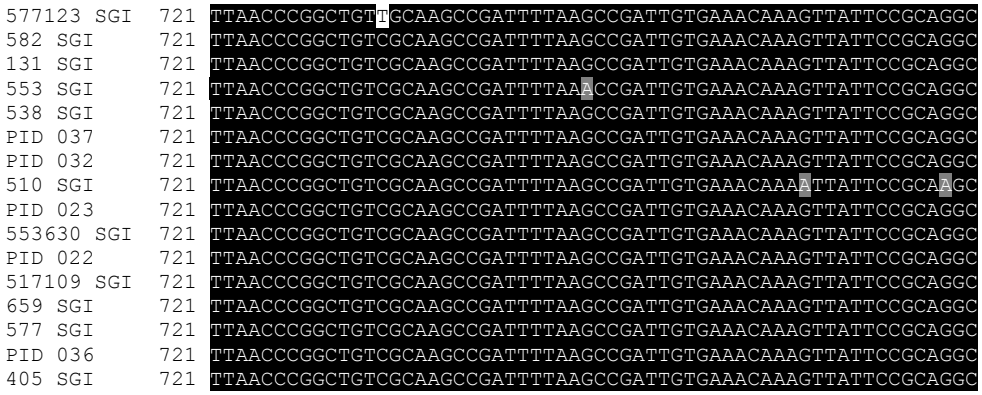
Shown are the *lpt3* nucleotide alignments from isolates taken from patients with UG. Shaded regions indicate areas of conservation. Gray and white boxes indicate where nucleotide differences exist.




577123 SGI 181 TTGATTGCGGTGTTTTTCGCGTTCAGCATTATTGCCAACAATGTGCATTACGCGGTTTAT
582 SGI 181 TTGATTGCGGTGTTTTTCGCGTTCAGCATTATTGCCAACAATGTGCATTACGCGGTTTAT
131 SGI 181 TTGATTGCGGTGTTTTTCGCGTTCATTATTATTGCCAACAATGTGCATTACGCGGTTTAT
553 SGI 181 TTGATTGCGGTGTTTTTCGCGTTCAGCATTATTGCCAACAATGTGCATTACGCGGTTTAT
538 SGI 181 TTGATTGCGGTGTTTTTCGCGTTCAGCATTATTGCCAACAATGTGCATTACGCGGTTTAT
PID 037 181 TTGATTGCGGTGTTTTTCGCGTTCAGCATTATTGCCAACAATGTGCATTACGCGGTTTAT
PID 032 181 TTGATTGCGGTGTTTTTCGCGTTCAGCATTATTGCCAACAATGTGCATTACGCGGTTTAT
510 SGI 181 CTGATTGCGGTGTTTTTCGCGTTCAGCATCATTGCCAACAATGTGCATTACGCGGTTTAT
PID 023 181 CTGATTGCGGTGTTTTTCGCGTTCAGCATCATTGCCAACAATGTGCATTACGCGGTTTAT
553630 SGI 181 TTGATTGCGGTGTTTTTCGCGTTCAGCATTATTGCCAACAATGTGCATTACGCGGTTTAT
PID 022 181 CTGATTGCGGTGTTTTTCGCGTTCAGCATCATTGCCAACAATGTGCATTACGCGGTTTAT
517109 SGI 181 TTGATTGCGGTGTTTTTCGCGTTCAGCATTATTGCCAACAATGTGCATTACGCGGTTTAT
659 SGI 181 TTGATTGCGGTGTTTTTCGCGTTCAGCATTATTGCCAACAATGTGCATTACGCGGTTTAT
577 SGI 181 TTGATTGCGGTGTTTTTCGCGTTCAGCATTATTGCCAACAATGTGCATTACGCGGTTTAT
PID 036 181 CTGATTGCGGTGTTTTTCGCGTTCAGCATCATTGCCAACAATGTGCATTACGCGGTTTAT
405 SGI 181 CTGATTGCGGTGTTTTTCGCGTTCAGCATCATTGCCAACAATGTGCATTACGCGGTTTAT
577123 SGI 241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
582 SGI 241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
131 SGI 241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
553 SGI 241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
538 SGI 241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
PID 037 241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
PID 032 241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
510 SGI 241 CAAAGCTGGATGACGGGCATCAATTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
PID 023 241 CAAAGCTGGATGACGGGCATCAATTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
553630 SGI 241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
PID 022 241 CAAAGCTGGATGACGGGCATCAATTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
517109 SGI 241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
659 SGI 241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
577 SGI 241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
PID 036 241 CAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
405 SGI 241 CAAAGCTGGATGACGGGCATCAATTATTGGCTGATGCTGAAAGAGGTTACCGAAGTCGGC
577123 SGI 301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
582 SGI 301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
131 SGI 301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
553 SGI 301 AGCGCGGGCGCGTCAATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
538 SGI 301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
PID 037 301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
PID 032 301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
510 SGI 301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
PID 023 301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
553630 SGI 301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
PID 022 301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
517109 SGI 301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
659 SGI 301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
577 SGI 301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
PID 036 301 AGCGCGGGCGCGTCGATGTTGGATAAAATTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
405 SGI 301 AGCGCGGGCGCGTCGATGTTGGATAAGTTGTGGCTGCCTGCTTTGTGGGGCGTGGCGGAA
577123 SGI 361 GTCATGTTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATA
582 SGI 361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
131 SGI 361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
553 SGI 361 GTCATGTTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATA
538 SGI 361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
PID 037 361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
PID 032 361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
510 SGI 361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
PID 023 361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
553630 SGI 361 GTCATGTTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATA
PID 022 361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
517109 SGI 361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
659 SGI 361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
577 SGI 361 GTCATGTTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATA
PID 036 361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC
405 SGI 361 GTCGTGCTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCATTTTTCTGCCGATATC

```
577123 SGI  421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
582 SGI      421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
131 SGI      421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
553 SGI      421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
538 SGI      421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
PID 037      421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
PID 032      421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
510 SGI      421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
PID 023      421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
553630 SGI   421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
PID 022      421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
517109 SGI   421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
659 SGI      421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
577 SGI      421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
PID 036      421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC
405 SGI      421 CTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTTCGTTCGACACGAAACAAGAGCAC

577123 SGI  481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
582 SGI      481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
131 SGI      481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
553 SGI      481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
538 SGI      481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
PID 037      481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
PID 032      481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
510 SGI      481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
PID 023      481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
553630 SGI   481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
PID 022      481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
517109 SGI   481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
659 SGI      481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
577 SGI      481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
PID 036      481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT
405 SGI      481 GGTATTTCGCCCAAACCGACATACAGCCGCATCAAAGCCAATTATTTCAGCTTCGGTTAT

577123 SGI  541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
582 SGI      541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
131 SGI      541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
553 SGI      541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
538 SGI      541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
PID 037      541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
PID 032      541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
510 SGI      541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
PID 023      541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
553630 SGI   541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
PID 022      541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
517109 SGI   541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
659 SGI      541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
577 SGI      541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
PID 036      541 TTTGTCGGGCGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA
405 SGI      541 TTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATTTAAGCAAGATCCCTGTGTTCAAA

577123 SGI  601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
582 SGI      601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
131 SGI      601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
553 SGI      601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
538 SGI      601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
PID 037      601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
PID 032      601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
510 SGI      601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
PID 023      601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
553630 SGI   601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
PID 022      601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
517109 SGI   601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
659 SGI      601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
577 SGI      601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
PID 036      601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
405 SGI      601 CAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTCAAAATATCGTCCTGATTATGGGC
```


577123 SGI 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAAATTCGCCGTTT
582 SGI 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
131 SGI 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
553 SGI 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCTCCGTTT
538 SGI 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
PID 037 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
PID 032 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
510 SGI 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
PID 023 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
553630 SGI 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
PID 022 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
517109 SGI 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
659 SGI 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
577 SGI 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
PID 036 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT
405 SGI 661 GAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTGGTTACGGGCGCGAAACTTCGCCGTTT

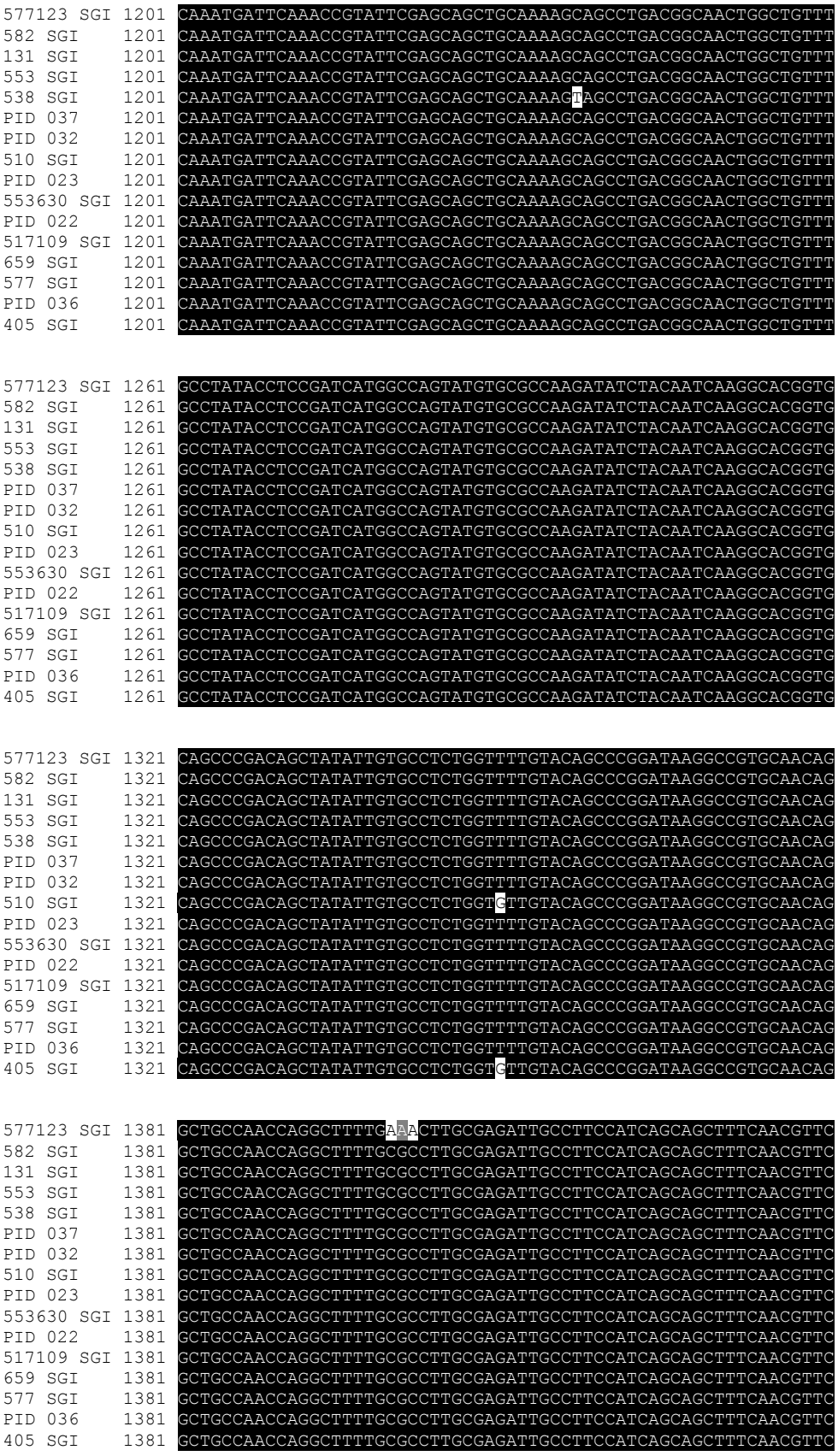

577123 SGI 721 TTAACCCGGCTGTTGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
582 SGI 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
131 SGI 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
553 SGI 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAACCGATTGTGAAACAAAGTTATTCCGCAGGC
538 SGI 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
PID 037 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
PID 032 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
510 SGI 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAATTATTCCGCAAGC
PID 023 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
553630 SGI 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
PID 022 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
517109 SGI 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
659 SGI 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
577 SGI 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
PID 036 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC
405 SGI 721 TTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGATTGTGAAACAAAGTTATTCCGCAGGC

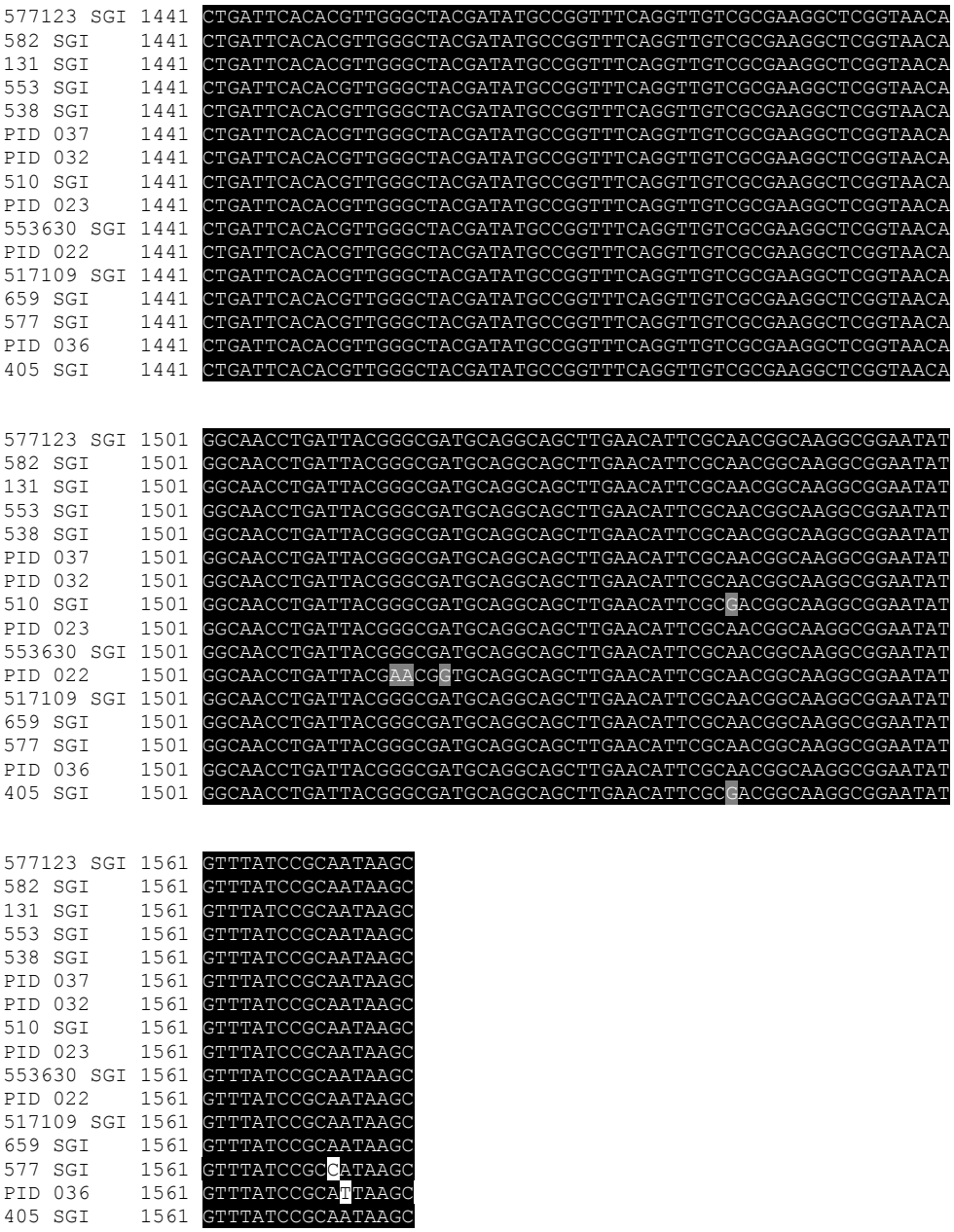

577123 SGI 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
582 SGI 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
131 SGI 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
553 SGI 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
538 SGI 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
PID 037 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
PID 032 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
510 SGI 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
PID 023 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
553630 SGI 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
PID 022 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
517109 SGI 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
659 SGI 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
577 SGI 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG
PID 036 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAAACGCTTG
405 SGI 781 TTTATGACGGCAGTATCCCTGCCCAGTTTCTTTAACGTCATACCGCACGCCAACGGCTTG


577123 SGI 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
582 SGI 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
131 SGI 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
553 SGI 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
538 SGI 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
PID 037 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
PID 032 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
510 SGI 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
PID 023 841 TAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
553630 SGI 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
PID 022 841 AAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
517109 SGI 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
659 SGI 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
577 SGI 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
PID 036 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA
405 SGI 841 GAACAAATCAGCGGCGGCGATACCAATATGTTCCGCCTCGCCAAAGAGCAGGGCTATGAA

```
577123 SGI  901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
582 SGI      901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
131 SGI      901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
553 SGI      901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
538 SGI      901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
PID 037      901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
PID 032      901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
510 SGI      901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
PID 023      901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
553630 SGI   901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
PID 022      901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
517109 SGI   901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
659 SGI      901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
577 SGI      901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
PID 036      901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG
405 SGI      901 ACGTATTTTTACAGTGCCCAGGCTGAAAACCAAATGGCAATTTTGAACTTAATCGGTAAG

577123 SGI  961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
582 SGI      961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
131 SGI      961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
553 SGI      961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
538 SGI      961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
PID 037      961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
PID 032      961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
510 SGI      961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
PID 023      961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
553630 SGI   961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
PID 022      961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
517109 SGI   961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
659 SGI      961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
577 SGI      961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
PID 036      961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG
405 SGI      961 AAATGGATAGACCATCTGATTCAGCCGACGCAACTTGGCTACGGCAACGGCGACAATATG

577123 SGI 1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
582 SGI    1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
131 SGI    1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
553 SGI    1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
538 SGI    1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
PID 037    1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
PID 032    1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
510 SGI    1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
PID 023    1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
553630 SGI 1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
PID 022    1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
517109 SGI 1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
659 SGI    1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
577 SGI    1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
PID 036    1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT
405 SGI    1021 CCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAATCAATTTGCAGCAGGGCAGGCATTTT

577123 SGI 1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
582 SGI    1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
131 SGI    1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
553 SGI    1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
538 SGI    1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
PID 037    1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
PID 032    1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
510 SGI    1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
PID 023    1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
553630 SGI 1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
PID 022    1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
517109 SGI 1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
659 SGI    1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
577 SGI    1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
PID 036    1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
405 SGI    1081 ATCGTGTTGCACCAACGCGGTTCGCACGCCCCATACGGCGCATTGTTGCAGCCTCAAGAT
```

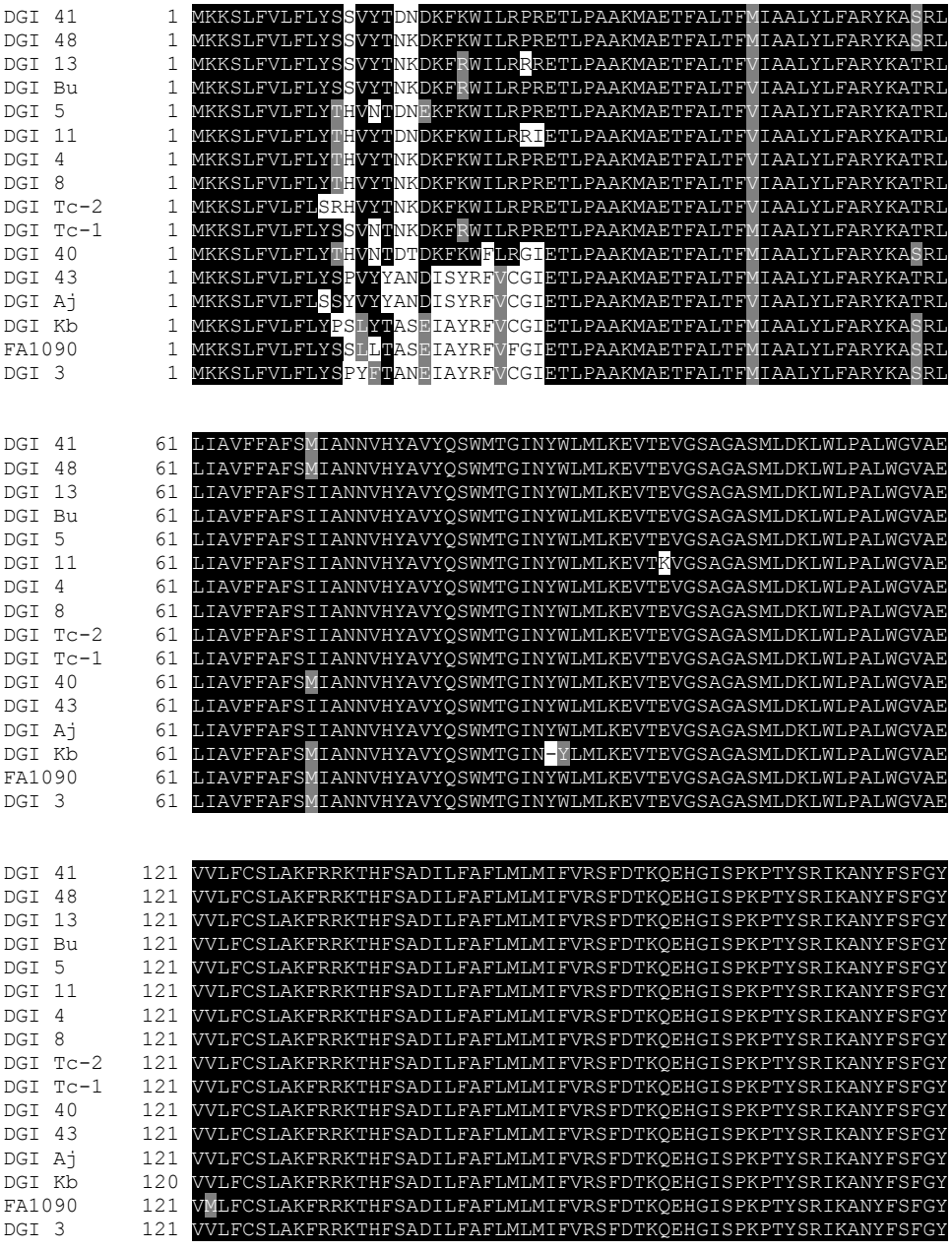

577123 SGI 1141 AAAGTATTCGGCGAAGCCGATATTGTGGATAAGTAAGACAACACCATCCACAAAACCGAC
582 SGI 1141 AAAGTATTCGGCGAAGCCGATATTGCGGATAAGTACGACAACACCATCCACAAAACCGAC
131 SGI 1141 AAAGTATTCGGCGAAGCCGATATTGCGGATAAGTACGACAACACCATCCACAAAACCGAC
553 SGI 1141 AAAGTATTCGGCGAAGCCGATATTGTGGATAAGTACGACAACACCATCCACAAAACCGAC
538 SGI 1141 AAAGTATTCGGCGAAGCCGATATTGCGGATAAGTACGACAACACCATCCACAAAACCGAC
PID 037 1141 AAAGTATTCGGCGAAGCCGATATTGCGGATAAGTACGACAACACCATCCACAAAACCGAC
PID 032 1141 AAAGTATTCGGCGAAGCCGATATTGCGGATAAGTACGACAACACCATCCACAAAACCGAC
510 SGI 1141 AAAGTATTCGGCGAAGCCGATATTGCGGATAAGTACGACAACACCATCCACAAAACCGAC
PID 023 1141 AAAGTATTCGGCGAAGCCGATATTGTGGATAAGTACGACAACACCATCCACAAAACCGAC
553630 SGI 1141 AAAGTATTCGGCGAAGCCGATATTGTGGATAAGTACGACAACACCATCCACAAAACCGAC
PID 022 1141 AAAGTATTCGGCGAAGCCGATATTGTGGATAAGTACGACAACACCATCCACAAAACCGAC
517109 SGI 1141 AAAGTATTCGGCGAAGCCGATATTGCGGATAAGTACGACAACACCATCCACAAAACCGAC
659 SGI 1141 AAAGTATTCGGCGAAGCCGATATTGCGGATAAGTACGACAACACCATCCACAAAACCGAC
577 SGI 1141 AAAGTATTCGGCGAAGCCGATATTGTGGATAAGTACGACAACACCATCCACAAAACCGAC
PID 036 1141 AAAGTATTCGGCGAAGCCGATATTGCGGATAAGTACGACAACACCATCCACAAAACCGAC
405 SGI 1141 AAAGTATTCGGCGAAGCCGATATTGCGGATAAGTACGACAACACCATCCACAAAACCGAC


577123 SGI 1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
582 SGI 1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
131 SGI 1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
553 SGI 1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
538 SGI 1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGTAGCCTGACGGCAACTGGCTGTTT
PID 037 1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
PID 032 1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
510 SGI 1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
PID 023 1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
553630 SGI 1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
PID 022 1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
517109 SGI 1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
659 SGI 1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
577 SGI 1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
PID 036 1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
405 SGI 1201 CAAATGATTCAAACCGTATTCGAGCAGCTGCAAAAGCAGCCTGACGGCAACTGGCTGTTT
577123 SGI 1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
582 SGI 1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
131 SGI 1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
553 SGI 1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
538 SGI 1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
PID 037 1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
PID 032 1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
510 SGI 1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
PID 023 1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
553630 SGI 1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
PID 022 1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
517109 SGI 1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
659 SGI 1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
577 SGI 1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
PID 036 1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
405 SGI 1261 GCCTATACCTCCGATCATGGCCAGTATGTGCGCCAAGATATCTACAATCAAGGCACGGTG
577123 SGI 1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
582 SGI 1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
131 SGI 1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
553 SGI 1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
538 SGI 1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
PID 037 1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
PID 032 1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
510 SGI 1321 CAGCCCGACAGCTATATTGTGCCTCTGGTGTTGTACAGCCCGGATAAGGCCGTGCAACAG
PID 023 1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
553630 SGI 1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
PID 022 1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
517109 SGI 1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
659 SGI 1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
577 SGI 1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
PID 036 1321 CAGCCCGACAGCTATATTGTGCCTCTGGTTTTGTACAGCCCGGATAAGGCCGTGCAACAG
405 SGI 1321 CAGCCCGACAGCTATATTGTGCCTCTGGTGTTGTACAGCCCGGATAAGGCCGTGCAACAG
577123 SGI 1381 GCTGCCAACCAGGCTTTTGAAACTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
582 SGI 1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
131 SGI 1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
553 SGI 1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
538 SGI 1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
PID 037 1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
PID 032 1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
510 SGI 1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
PID 023 1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
553630 SGI 1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
PID 022 1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
517109 SGI 1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
659 SGI 1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
577 SGI 1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
PID 036 1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC
405 SGI 1381 GCTGCCAACCAGGCTTTTGCGCCTTGCGAGATTGCCTTCCATCAGCAGCTTTCAACGTTC

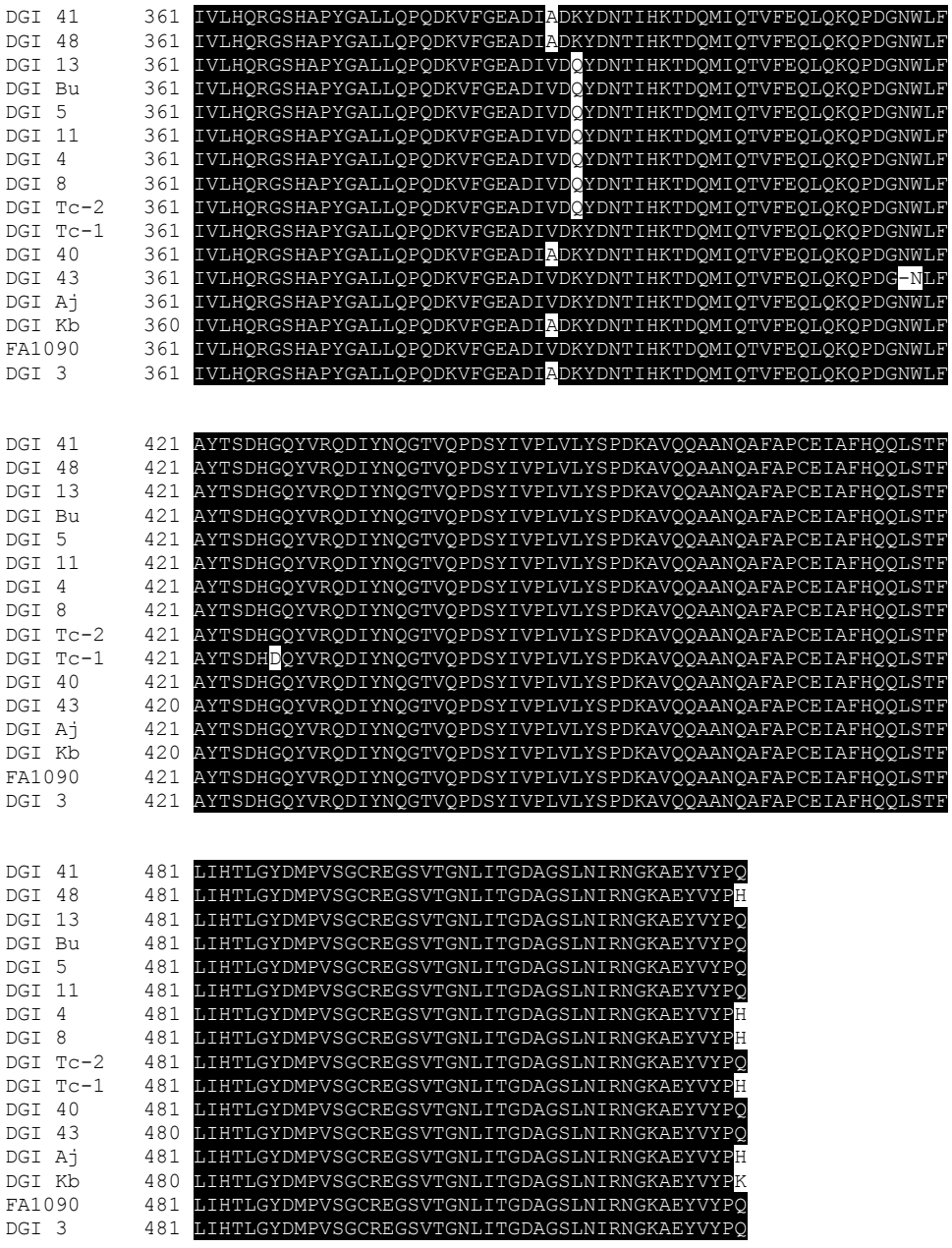

577123 SGI 1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
582 SGI 1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
131 SGI 1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
553 SGI 1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
538 SGI 1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
PID 037 1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
PID 032 1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
510 SGI 1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
PID 023 1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
553630 SGI 1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
PID 022 1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
517109 SGI 1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
659 SGI 1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
577 SGI 1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
PID 036 1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
405 SGI 1441 CTGATTCACACGTTGGGCTACGATATGCCGGTTTCAGGTTGTCGCGAAGGCTCGGTAACA
577123 SGI 1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
582 SGI 1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
131 SGI 1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
553 SGI 1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
538 SGI 1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
PID 037 1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
PID 032 1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
510 SGI 1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCGACGGCAAGGCGGAATAT
PID 023 1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
553630 SGI 1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
PID 022 1501 GGCAACCTGATTACGAACGGTGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
517109 SGI 1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
659 SGI 1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
577 SGI 1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
PID 036 1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCAACGGCAAGGCGGAATAT
405 SGI 1501 GGCAACCTGATTACGGGCGATGCAGGCAGCTTGAACATTCGCGACGGCAAGGCGGAATAT
577123 SGI 1561 GTTTATCCGCAATAAGC
582 SGI 1561 GTTTATCCGCAATAAGC
131 SGI 1561 GTTTATCCGCAATAAGC
553 SGI 1561 GTTTATCCGCAATAAGC
538 SGI 1561 GTTTATCCGCAATAAGC
PID 037 1561 GTTTATCCGCAATAAGC
PID 032 1561 GTTTATCCGCAATAAGC
510 SGI 1561 GTTTATCCGCAATAAGC
PID 023 1561 GTTTATCCGCAATAAGC
553630 SGI 1561 GTTTATCCGCAATAAGC
PID 022 1561 GTTTATCCGCAATAAGC
517109 SGI 1561 GTTTATCCGCAATAAGC
659 SGI 1561 GTTTATCCGCAATAAGC
577 SGI 1561 GTTTATCCGCCATAAGC
PID 036 1561 GTTTATCCGCATTAAGC
405 SGI 1561 GTTTATCCGCAATAAGC

Appendix B. Amino acid sequences of Lpt3 from *N. gonorrhoeae*.

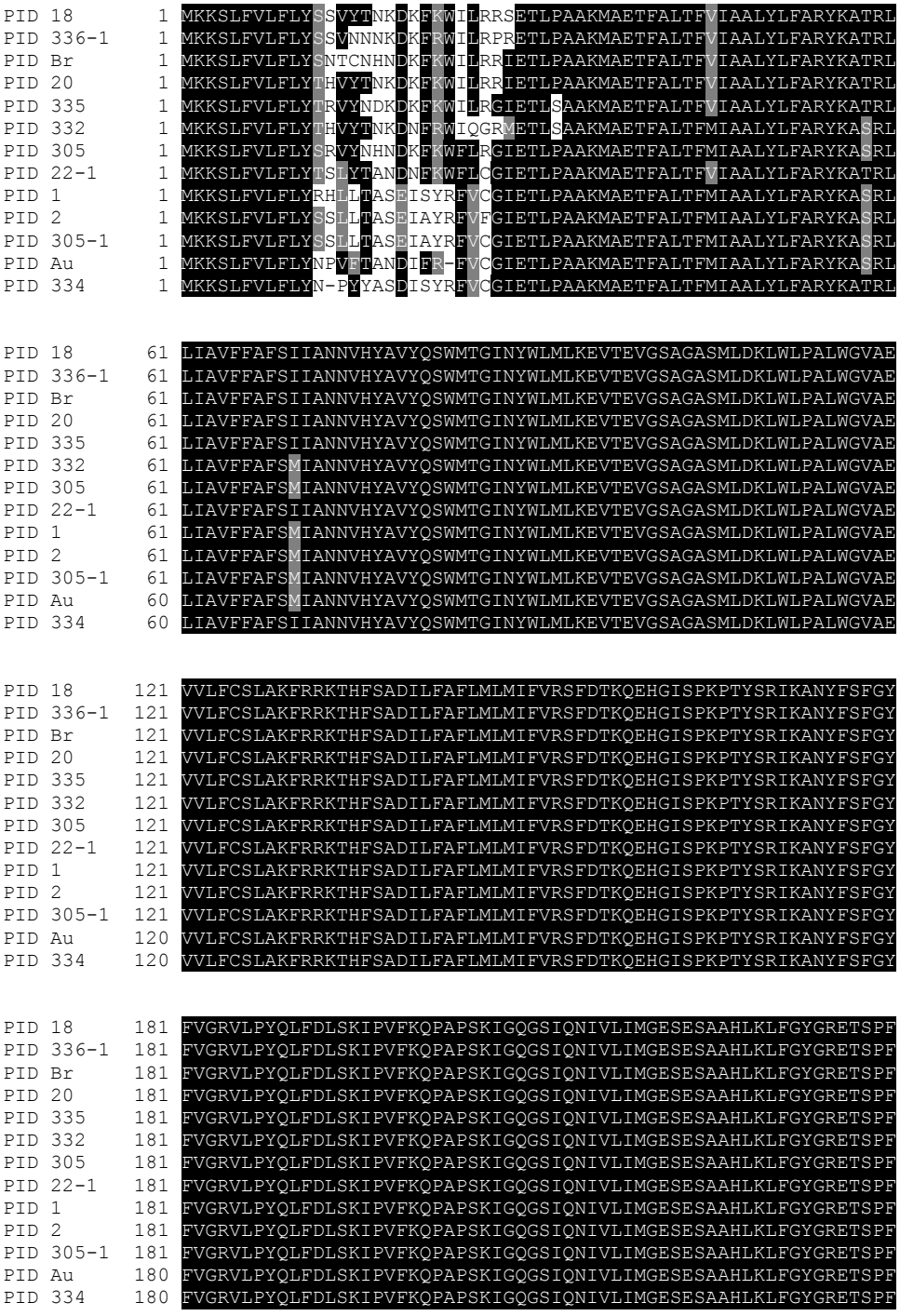
Shown are the Lpt3 amino acid alignments from isolates taken from patients with DGI. Shaded regions indicate areas of conservation. Gray boxes indicate where minor amino acid differences exist, and white areas show sites of more significant variation.



```
DGI 41     181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
DGI 48     181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
DGI 13     181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
DGI Bu     181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
DGI 5      181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
DGI 11     181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
DGI 4      181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
DGI 8      181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
DGI Tc-2   181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
DGI Tc-1   181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
DGI 40     181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
DGI 43     181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
DGI Aj     181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
DGI Kb     180 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
FA1090     181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
DGI 3      181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF

DGI 41     241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
DGI 48     241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
DGI 13     241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGFVQISGGDTNMFRLAKEQGYE
DGI Bu     241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
DGI 5      241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
DGI 11     241 LTRLSQADFKPIVKQSYSAGFMTAVSLSSFFNVIPHANGLVQISGGDTNMFRLAKEQGYE
DGI 4      241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
DGI 8      241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
DGI Tc-2   241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLYQISGGDTNMFRLAKEQGYE
DGI Tc-1   241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
DGI 40     241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
DGI 43     241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
DGI Aj     241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
DGI Kb     240 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
FA1090     241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
DGI 3      241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE

DGI 41     301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
DGI 48     301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
DGI 13     301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
DGI Bu     301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
DGI 5      301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
DGI 11     301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
DGI 4      301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
DGI 8      301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
DGI Tc-2   301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
DGI Tc-1   301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
DGI 40     301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
DGI 43     301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
DGI Aj     301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
DGI Kb     300 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
FA1090     301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
DGI 3      301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
```


DGI 41 361 IVLHQRGSHAPYGALLQPQDKVFGEADIADKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
DGI 48 361 IVLHQRGSHAPYGALLQPQDKVFGEADIADKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
DGI 13 361 IVLHQRGSHAPYGALLQPQDKVFGEADIVDQYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
DGI Bu 361 IVLHQRGSHAPYGALLQPQDKVFGEADIVDQYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
DGI 5 361 IVLHQRGSHAPYGALLQPQDKVFGEADIVDQYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
DGI 11 361 IVLHQRGSHAPYGALLQPQDKVFGEADIVDQYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
DGI 4 361 IVLHQRGSHAPYGALLQPQDKVFGEADIVDQYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
DGI 8 361 IVLHQRGSHAPYGALLQPQDKVFGEADIVDQYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
DGI Tc-2 361 IVLHQRGSHAPYGALLQPQDKVFGEADIVDQYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
DGI Tc-1 361 IVLHQRGSHAPYGALLQPQDKVFGEADIVDKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
DGI 40 361 IVLHQRGSHAPYGALLQPQDKVFGEADIADKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
DGI 43 361 IVLHQRGSHAPYGALLQPQDKVFGEADIVDKYDNTIHKTDQMIQTVFEQLQKQPDG-NLF
DGI Aj 361 IVLHQRGSHAPYGALLQPQDKVFGEADIVDKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
DGI Kb 360 IVLHQRGSHAPYGALLQPQDKVFGEADIADKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
FA1090 361 IVLHQRGSHAPYGALLQPQDKVFGEADIVDKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
DGI 3 361 IVLHQRGSHAPYGALLQPQDKVFGEADIADKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
DGI 41 421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
DGI 48 421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
DGI 13 421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
DGI Bu 421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
DGI 5 421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
DGI 11 421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
DGI 4 421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
DGI 8 421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
DGI Tc-2 421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
DGI Tc-1 421 AYTSDHDQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
DGI 40 421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
DGI 43 420 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
DGI Aj 421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
DGI Kb 420 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
FA1090 421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
DGI 3 421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
DGI 41 481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
DGI 48 481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPH
DGI 13 481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
DGI Bu 481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
DGI 5 481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
DGI 11 481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
DGI 4 481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPH
DGI 8 481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPH
DGI Tc-2 481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
DGI Tc-1 481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPH
DGI 40 481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
DGI 43 480 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
DGI Aj 481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPH
DGI Kb 480 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPK
FA1090 481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
DGI 3 481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ

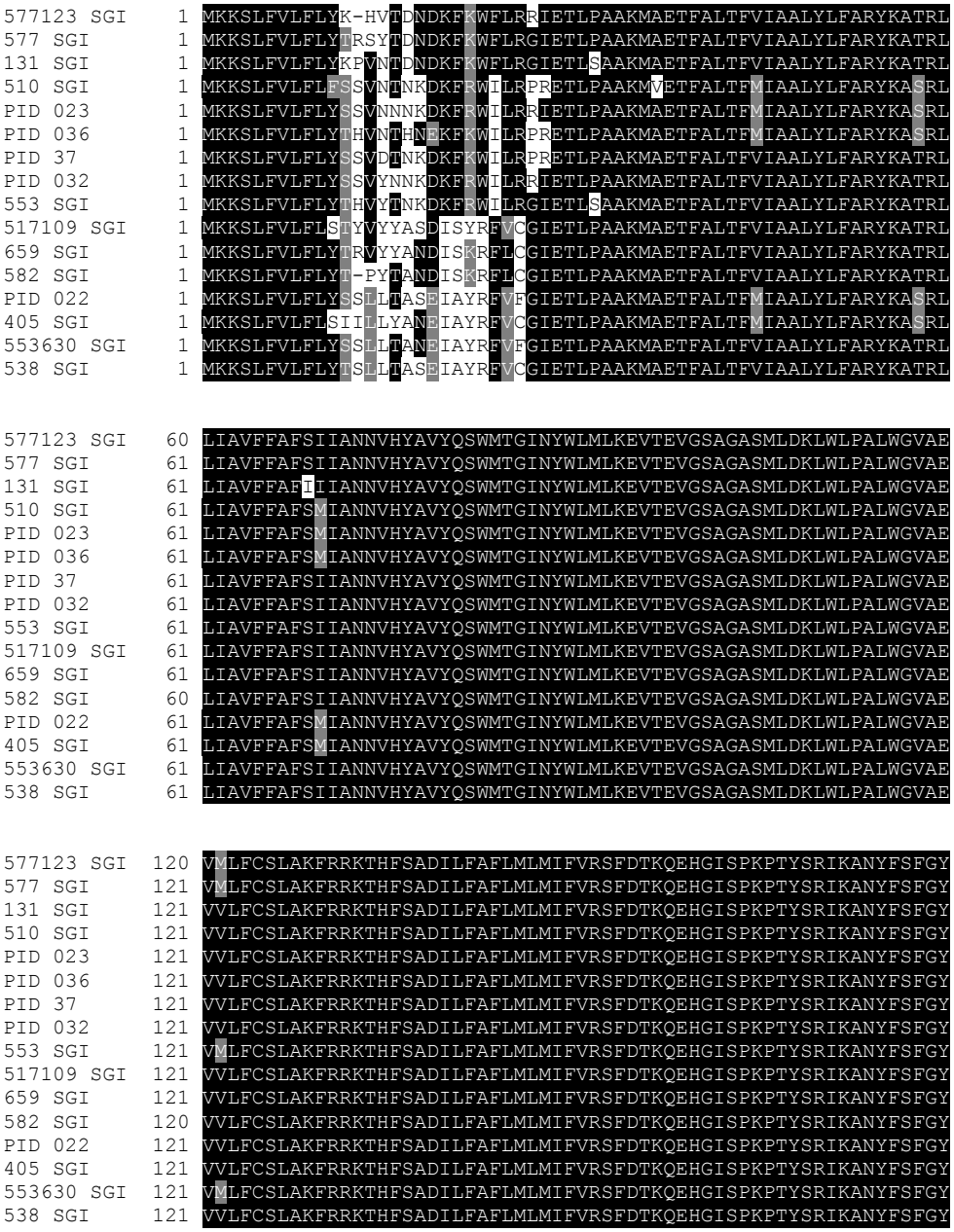
Shown are the Lpt3 amino acid alignments from isolates taken from patients with PID. Shaded regions indicate areas of conservation. Gray boxes indicate where minor amino acid differences exist, and white areas show sites of more significant variation.

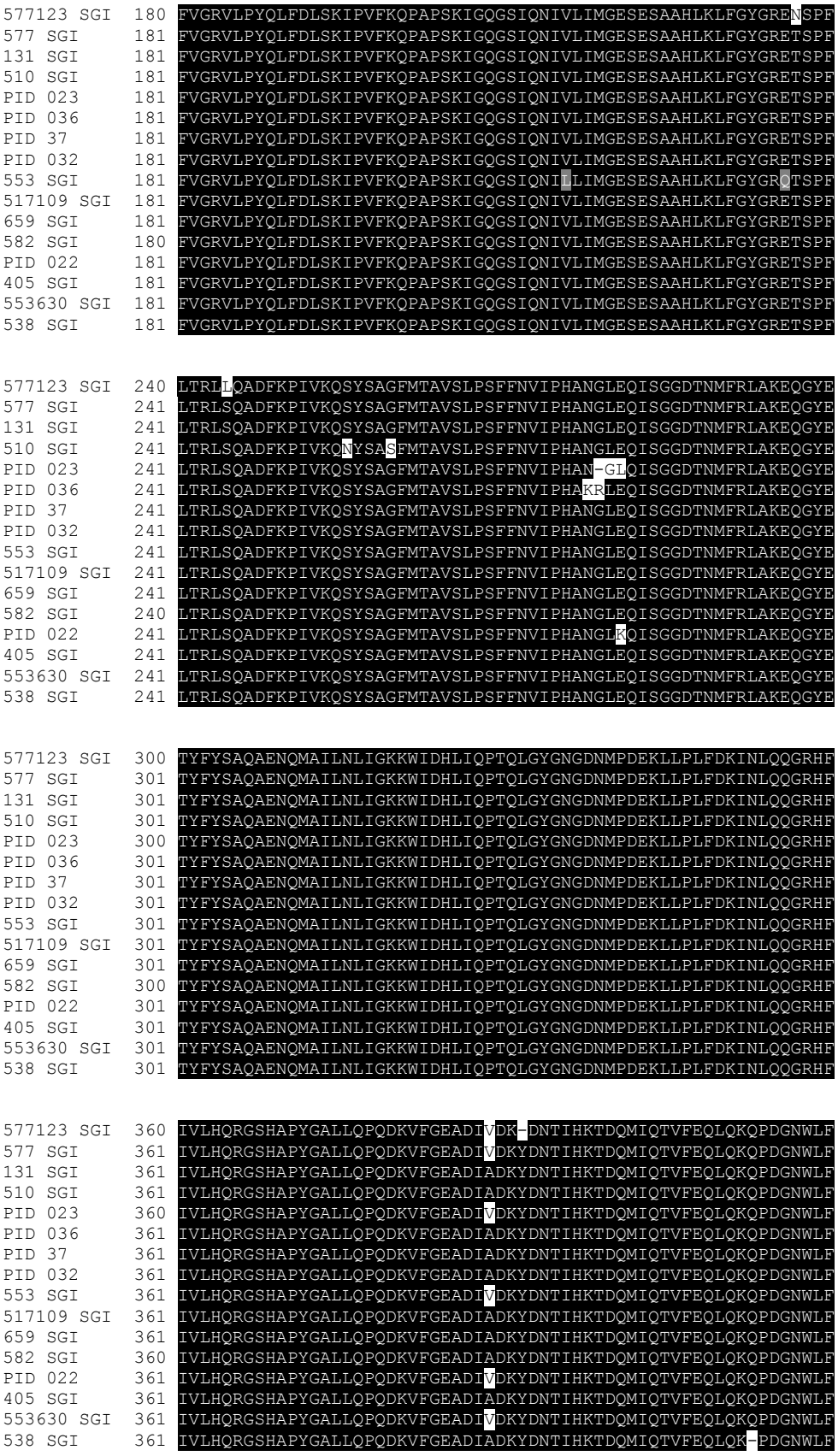


```
PID 18     241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGFLQSSGGETNMFRLAKEQGYE
PID 336-1  241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
PID Br     241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
PID 20     241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
PID 335    241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
PID 332    241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
PID 305    241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
PID 22-1   241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
PID 1      241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
PID 2      241 LTRLSQADFKPICKQSYPAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
PID 305-1  241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
PID Au     240 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
PID 334    240 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE

PID 18     301 TYFYTFQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
PID 336-1  301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
PID Br     301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
PID 20     301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
PID 335    301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
PID 332    301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
PID 305    301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
PID 22-1   301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
PID 1      301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
PID 2      301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
PID 305-1  301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
PID Au     300 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
PID 334    300 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF

PID 18     361 IVLHQRGSHAPYGALLQPQDKVFGEADIVDKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
PID 336-1  361 IVLHQRGSHAPYGALLQPQDKVFGEADIVDKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
PID Br     361 IVLHQRGSHAPYGALLQPQDKVFGEADIVDKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
PID 20     361 IVLHQRGSHAPYGALLQPQDKVFGEADIVDKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
PID 335    361 IVLHQRGSHAPYGALLQPQDKVFGEADIVDKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
PID 332    361 IVLHQRGSHAPYGALLQPQDKVFGEADIVDKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
PID 305    361 IVLHQRGSHAPYGALLQPQDKVFGEADIVDKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
PID 22-1   361 IVLHQRGSHAPYGALLQPQDKVFGEADIVDKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
PID 1      361 IVLHQRGSHAPYGALLQPQDKVFGEADIADKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
PID 2      361 IVLHQRGSHAPYGALLQPQDKVFGEADIADKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
PID 305-1  361 IVLHQRGSHAPYGALLQPQDKVFGEADIVDKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
PID Au     360 IVLHQRGSHAPYGALLQPQDKVFGEADIADKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
PID 334    360 IVLHQRGSHAPYGALLQPQDKVFGEADIVDKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF

PID 18     421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
PID 336-1  421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
PID Br     421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
PID 20     421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
PID 335    421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
PID 332    421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
PID 305    421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
PID 22-1   421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
PID 1      421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
PID 2      421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
PID 305-1  421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
PID Au     420 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
PID 334    420 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
```

```
PID 18      481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
PID 336-1   481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
PID Br      481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
PID 20      481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
PID 335     481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
PID 332     481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
PID 305     481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
PID 22-1    481 LIHTLGYDMPVSGCRESSVTGNLITGDAGSLNIRNGKAEYVYPQ
PID 1       481 LIHTLGYDMPVSGCREGSVTGNLITGDASSLNIRNGKAEYVYPH
PID 2       481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPH
PID 305-1   481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
PID Au      480 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
PID 334     480 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
```

Shown are the Lpt3 amino acid alignments from isolates taken from patients with UG. Shaded regions indicate areas of conservation. Gray boxes indicate where minor amino acid differences exist, and white areas show sites of more significant variation.




577123 SGI 180 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRENSPF
577 SGI 181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
131 SGI 181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
510 SGI 181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
PID 023 181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
PID 036 181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
PID 37 181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
PID 032 181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
553 SGI 181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNILLIMGESESAAHLKLFGYGRQTSPF
517109 SGI 181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
659 SGI 181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
582 SGI 180 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
PID 022 181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
405 SGI 181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
553630 SGI 181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
538 SGI 181 FVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLIMGESESAAHLKLFGYGRETSPF
577123 SGI 240 LTRLLQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
577 SGI 241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
131 SGI 241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
510 SGI 241 LTRLSQADFKPIVKQNYSASFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
PID 023 241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHAN-GLQISGGDTNMFRLAKEQGYE
PID 036 241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHAKRLEQISGGDTNMFRLAKEQGYE
PID 37 241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
PID 032 241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
553 SGI 241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
517109 SGI 241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
659 SGI 241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
582 SGI 240 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
PID 022 241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLKQISGGDTNMFRLAKEQGYE
405 SGI 241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
553630 SGI 241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
538 SGI 241 LTRLSQADFKPIVKQSYSAGFMTAVSLPSFFNVIPHANGLEQISGGDTNMFRLAKEQGYE
577123 SGI 300 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
577 SGI 301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
131 SGI 301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
510 SGI 301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
PID 023 300 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
PID 036 301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
PID 37 301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
PID 032 301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
553 SGI 301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
517109 SGI 301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
659 SGI 301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
582 SGI 300 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
PID 022 301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
405 SGI 301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
553630 SGI 301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
538 SGI 301 TYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGYGNGDNMPDEKLLPLFDKINLQQGRHF
577123 SGI 360 IVLHQRGSHAPYGALLQPQDKVFGEADIVDK-DNTIHKTDQMIQTVFEQLQKQPDGNWLF
577 SGI 361 IVLHQRGSHAPYGALLQPQDKVFGEADIVDKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
131 SGI 361 IVLHQRGSHAPYGALLQPQDKVFGEADIADKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
510 SGI 361 IVLHQRGSHAPYGALLQPQDKVFGEADIADKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
PID 023 360 IVLHQRGSHAPYGALLQPQDKVFGEADIVDKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
PID 036 361 IVLHQRGSHAPYGALLQPQDKVFGEADIADKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
PID 37 361 IVLHQRGSHAPYGALLQPQDKVFGEADIADKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
PID 032 361 IVLHQRGSHAPYGALLQPQDKVFGEADIADKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
553 SGI 361 IVLHQRGSHAPYGALLQPQDKVFGEADIVDKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
517109 SGI 361 IVLHQRGSHAPYGALLQPQDKVFGEADIADKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
659 SGI 361 IVLHQRGSHAPYGALLQPQDKVFGEADIADKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
582 SGI 360 IVLHQRGSHAPYGALLQPQDKVFGEADIADKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
PID 022 361 IVLHQRGSHAPYGALLQPQDKVFGEADIVDKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
405 SGI 361 IVLHQRGSHAPYGALLQPQDKVFGEADIADKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
553630 SGI 361 IVLHQRGSHAPYGALLQPQDKVFGEADIVDKYDNTIHKTDQMIQTVFEQLQKQPDGNWLF
538 SGI 361 IVLHQRGSHAPYGALLQPQDKVFGEADIADKYDNTIHKTDQMIQTVFEQLQK-PDGNWLF

```
577123 SGI  419 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFETCEIAFHQQLSTF
577 SGI      421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
131 SGI      421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
510 SGI      421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
PID 023      420 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
PID 036      421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
PID 37       421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
PID 032      421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
553 SGI      421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
517109 SGI  421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
659 SGI      421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
582 SGI      420 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
PID 022      421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
405 SGI      421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
553630 SGI  421 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
538 SGI      420 AYTSDHGQYVRQDIYNQGTVQPDSYIVPLVLYSPDKAVQQAANQAFAPCEIAFHQQLSTF
```

```
577123 SGI  479 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
577 SGI      481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPP
131 SGI      481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
510 SGI      481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRDGKAEYVYPQ
PID 023      480 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
PID 036      481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPH
PID 37       481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
PID 032      481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
553 SGI      481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
517109 SGI  481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
659 SGI      481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
582 SGI      480 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
PID 022      481 LIHTLGYDMPVSGCREGSVTGNLITNGAGSLNIRNGKAEYVYPQ
405 SGI      481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRDGKAEYVYPQ
553630 SGI  481 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
538 SGI      480 LIHTLGYDMPVSGCREGSVTGNLITGDAGSLNIRNGKAEYVYPQ
```

Appendix C: Outline of sample preparation prior to GC-MS Analysis

- Sample dried in 13X100 screw-cap test tube
- Note L-Rhamnose, L(-)Fucose, and GlcNAc were run concurrently as standards
- HCl (1N) in anhydrous MeOH (1.0mL) added and tube flushed w/ N_2 before incubation at 80°C for 16-24hrs
- Evaporate to dryness under N_2 @ 35-40°C then 2X more after adding a few drops of MeOH to make sure HCl is eliminated
- Add MeOH (200µL), pyridine (20µL) and acetic anhydride (20µL)
- Vortex and stand 20min @ room temperature
- Evaporate to dryness under N_2 @ 35-40°C then 2X more after adding a few drops of toluene to azeotrope the pyridine, acetic acid and excess acetic anhydride. The monosaccharide methyl glycosides are stable at this point and can be stored in sealed tube indefinitely @ -20°C for later derivitization.
- Add Tri-Sil (200µL), flush tube with N_2 and place @ 80°C for 20-30min.
- Cool rapidly to room temperature, and the excess reagent is evaporated carefully just to the point of dryness under N_2 @ 35-40°C. White residue is washed 2X w/ n-hexane (100µL) combining washes in a small teflon lined screw-cap vial for GC-MS analysis.

Appendix D: Outline of sample preparation prior to MALDI-MS and ESI-MS

- 2 X 500µg placed in glass culture tubes; 1µg β–cyclodextrin added as a standard
- 200µL 1% acetic acid in water added to each tube
- Tubes placed on heating block for 1hr @ 100°C
- Tubes centrifuged and liquid layer was removed
- Liquid layer dried via vacuum centrifugation
- Reconstituted in water; MALDI-MS analysis performed
- Dried in Speed Vac.
- 1 sample derivatized with a pyrazole reducing-end protecting tag
- Both samples methylated and dried
- Reconstituted in 100µL MeOH
- Analyzed by MALDI and ESI-MS

References

1. **Aas, F. E., M. Wolfgang, S. Frye, S. Dunham, C. Lovold, and M. Koomey**. 2002. Competence for natural transformation in *Neisseria gonorrhoeae*: components of DNA binding and uptake linked to type IV pilus expression. Mol. Microbiol. **46**:749-760.

2. **Aho, E. L., A. M. Keating, and S. M. McGillivray**. 2000. A comparative analysis of pilin genes from pathogenic and nonpathogenic *Neisseria* species. Microb. Pathog. **28**:81-88.

3. **Andersen, S. R., G. Bjune, J. Lyngby, K. Bryn, and E. Jantzen**. 1995. Short-chain lipopolysaccharide mutants of serogroup B *Neisseria meningitidis* of potential value for production of outer membrane vesicle vaccines. Microb. Pathog. **19**:159-168.

4. **Andrews, T. D. and T. Gojobori**. 2004. Strong positive selection and recombination drive the antigenic variation of the PilE protein of the human pathogen *Neisseria meningitidis*. Genetics **166**:25-32.

5. **Apicella, M. A., M. A. Westerink, S. A. Morse, H. Schneider, P. A. Rice, and J. M. Griffiss**. 1986. Bactericidal antibody response of normal human serum to the lipooligosaccharide of *Neisseria gonorrhoeae*. J. Infect. Dis. **153**:520-526.

6. **Arking, D., Y. Tong, and D. C. Stein**. 2001. Analysis of lipooligosaccharide biosynthesis in the *Neisseriaceae*. J. Bacteriol. **183**:934-941.

7. **Arreaza, L., B. Alcala, C. Salcedo, F. L. De la, and J. A. Vazquez**. 2003. Dynamics of the penA gene in serogroup C meningococcal strains. J. Infect. Dis. **187**:1010-1014.

8. **Arreaza, L., C. Salcedo, B. Alcala, S. Berron, E. Martin, and J. A. Vazquez**. 2003. Antibiotic resistance of *Neisseria gonorrhoeae* in Spain: trends over the last two decades. J. Antimicrob. Chemother. **51**:153-156.

9. **Banerjee, A., R. Wang, S. L. Supernavage, S. K. Ghosh, J. Parker, N. F. Ganesh, P. G. Wang, S. Gulati, and P. A. Rice**. 2002. Implications of phase variation of a gene (pgtA) encoding a pilin galactosyl transferase in gonococcal pathogenesis. J. Exp. Med. **196**:147-162.

10. **Banerjee, A., R. Wang, S. N. Uljon, P. A. Rice, E. C. Gotschlich, and D. C. Stein**. 1998. Identification of the gene (lgtG) encoding the lipooligosaccharide beta chain synthesizing glucosyl transferase from *Neisseria gonorrhoeae*. Proc. Natl. Acad. Sci. U. S. A **95**:10872-10877.

11. **Berrington, A. W., Y. C. Tan, Y. Srikhanta, B. Kuipers, L. P. van der, I. R. Peak, and M. P. Jennings**. 2002. Phase variation in meningococcal lipooligosaccharide biosynthesis genes. FEMS Immunol. Med. Microbiol. **34**:267-275.

12. **Beutler, B. and F. Bazzoni**. 1998. TNF, apoptosis and autoimmunity: a common thread? Blood Cells Mol. Dis. **24**:216-230.

13. **Bille, E., J. R. Zahar, A. Perrin, S. Morelle, P. Kriz, K. A. Jolley, M. C. Maiden, C. Dervin, X. Nassif, and C. R. Tinsley**. 2005. A chromosomally integrated bacteriophage in invasive meningococci. J. Exp. Med. **201**:1905-1913.

14. **Birnboim, H. C. and J. Doly**. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513-1523.

15. **Bogdan, J. A., C. A. Minetti, and M. S. Blake**. 2002. A one-step method for genetic transformation of non-piliated *Neisseria meningitidis*. J. Microbiol. Methods **49**:97-101.

16. **Bonnah, R. A., J. Hoelter, L. Steeghs, C. A. Enns, M. So, and M. U. Muckenthaler**. 2005. Lipooligosaccharide-independent alteration of cellular homeostasis in *Neisseria meningitidis*-infected epithelial cells. Cell Microbiol. **7**:869-885.

17. **Brandtzaeg, P. and D. M. van**. 2002. Current concepts in the role of the host response in *Neisseria meningitidis* septic shock. Curr. Opin. Infect. Dis. **15**:247-252.

18. **Burch, C. L., R. J. Danaher, and D. C. Stein**. 1997. Antigenic variation in *Neisseria gonorrhoeae*: production of multiple lipooligosaccharides. J. Bacteriol. **179**:982-986.

19. **Chao, H. C., C. H. Chiu, Y. C. Huang, T. Y. Lin, and W. J. Su**. 1997. Endocarditis due to *Neisseria sicca*: report of one case. Zhonghua Min Guo. Xiao. Er. Ke. Yi. Xue. Hui. Za Zhi. **38**:229-231.

20. **Chaussee, M. S. and S. A. Hill**. 1998. Formation of single-stranded DNA during DNA transformation of *Neisseria gonorrhoeae*. J. Bacteriol. **180**:5117-5122.

21. **Cornelissen, C. N., J. E. Anderson, I. C. Boulton, and P. F. Sparling**. 2000. Antigenic and sequence diversity in gonococcal transferrin-binding protein A. Infect. Immun. **68**:4725-4735.

22. **Cox, A. D., J. Li, J. R. Brisson, E. R. Moxon, and J. C. Richards**. 2002. Structural analysis of the lipopolysaccharide from *Neisseria meningitidis* strain BZ157 galE: localisation of two phosphoethanolamine residues in the inner core oligosaccharide. Carbohydr. Res. **337**:1435-1444.

23. **Danaher, R. J., J. C. Levin, D. Arking, C. L. Burch, R. Sandlin, and D. C. Stein**. 1995. Genetic basis of *Neisseria gonorrhoeae* lipooligosaccharide antigenic variation. J. Bacteriol. **177**:7275-7279.

24. **De, C. C., A. Carannante, R. Lanzetta, R. Nunziata, V. Piscopo, and M. Parrilli**. 2004. Elucidation of two O-chain structures from the lipopolysaccharide fraction of *Agrobacterium tumefaciens* F/1. Carbohydr. Res. **339**:2451-2455.

25. **De, C. C., C. O. De, A. Molinaro, and M. Parrilli**. 2002. Structural determination of the O-chain polysaccharide from *Agrobacterium tumefaciens*, strain DSM 30205. Eur. J. Biochem. **269**:2885-2888.

26. **Deal, C. D. and H. C. Krivan**. 1990. Lacto- and ganglio-series glycolipids are adhesion receptors for *Neisseria gonorrhoeae*. J. Biol. Chem. **265**:12774-12777.

27. **Derrick, J. P., R. Urwin, J. Suker, I. M. Feavers, and M. C. Maiden**. 1999. Structural and evolutionary inference from molecular variation in *Neisseria* porins. Infect. Immun. **67**:2406-2413.

28. **Dillard, J. P. and H. S. Seifert**. 2001. A variable genetic island specific for *Neisseria gonorrhoeae* is involved in providing DNA for natural transformation and is found more often in disseminated infection isolates. Mol. Microbiol. **41**:263-277.

29. **Dinarello, C. A.** 1991. Inflammatory cytokines: interleukin-1 and tumor necrosis factor as effector molecules in autoimmune diseases. Curr. Opin. Immunol. **3**:941-948.

30. **Dowson, C. G., A. E. Jephcott, K. R. Gough, and B. G. Spratt**. 1989. Penicillin-binding protein 2 genes of non-beta-lactamase-producing, penicillin-resistant strains of *Neisseria gonorrhoeae*. Mol. Microbiol. **3**:35-41.

31. **Drazek, E. S., D. C. Stein, and C. D. Deal**. 1995. A mutation in the *Neisseria gonorrhoeae* rfaD homolog results in altered lipooligosaccharide expression. J. Bacteriol. **177**:2321-2327.

32. **Estabrook, M. M., J. M. Griffiss, and G. A. Jarvis**. 1997. Sialylation of *Neisseria meningitidis* lipooligosaccharide inhibits serum bactericidal activity by masking lacto-N-neotetraose. Infect. Immun. **65**:4436-4444.

33. **Facius, D. and T. F. Meyer**. 1993. A novel determinant (comA) essential for natural transformation competence in *Neisseria gonorrhoeae* and the effect of a comA defect on pilin variation. Mol. Microbiol. **10**:699-712.

34. **Feil, E., G. Carpenter, and B. G. Spratt**. 1995. Electrophoretic variation in adenylate kinase of *Neisseria meningitidis* is due to inter- and intraspecies recombination. Proc. Natl. Acad. Sci. U. S. A **92**:10535-10539.

35. **Feil, E. J., M. C. Maiden, M. Achtman, and B. G. Spratt**. 1999. The relative contributions of recombination and mutation to the divergence of clones of *Neisseria meningitidis*. Mol. Biol. Evol. **16**:1496-1502.

36. **Filiatrault, M. J., B. W. Gibson, B. Schilling, S. Sun, R. S. Munson, Jr., and A. A. Campagnari**. 2000. Construction and characterization of *Haemophilus ducreyi* lipooligosaccharide (LOS) mutants defective in expression of heptosyltransferase III and beta1,4-glucosyltransferase: identification of LOS glycoforms containing lactosamine repeats. Infect. Immun. **68**:3352-3361.

37. **Fitzpatrick, D. A., C. J. Creevey, and J. O. McInerney**. 2005. Evidence of Positive Darwinian Selection in Putative Meningococcal Vaccine Antigens. J. Mol. Evol.

38. **Fussenegger, M., D. Facius, J. Meier, and T. F. Meyer**. 1996. A novel peptidoglycan-linked lipoprotein (ComL) that functions in natural transformation competence of *Neisseria gonorrhoeae*. Mol. Microbiol. **19**:1095-1105.

39. **Fussenegger, M., A. F. Kahrs, D. Facius, and T. F. Meyer**. 1996. Tetrapac (tpc), a novel genotype of *Neisseria gonorrhoeae* affecting epithelial cell invasion, natural transformation competence and cell separation. Mol. Microbiol. **19**:1357-1372.

40. **Fussenegger, M., T. Rudel, R. Barten, R. Ryll, and T. F. Meyer**. 1997. Transformation competence and type-4 pilus biogenesis in *Neisseria gonorrhoeae*--a review. Gene **192**:125-134.

41. **Genco, C. A., S. A. Berish, C. Y. Chen, S. Morse, and D. L. Trees**. 1994. Genetic diversity of the iron-binding protein (Fbp) gene of the pathogenic and commensal *Neisseria*. FEMS Microbiol. Lett. **116**:123-129.

42. **Genco, C. A., C. Y. Chen, R. J. Arko, D. R. Kapczynski, and S. A. Morse**. 1991. Isolation and characterization of a mutant of *Neisseria gonorrhoeae* that is defective in the uptake of iron from transferrin and haemoglobin and is avirulent in mouse subcutaneous chambers. J. Gen. Microbiol. **137**:1313-1321.

43. **Genco, C. A. and P. J. Desai**. 1996. Iron acquisition in the pathogenic *Neisseria*. Trends Microbiol. **4**:179-184.

44. **Ghosh, S. K., J. Zhao, M. C. Philogene, A. Alzaharani, S. Rane, and A. Banerjee**. 2004. Pathogenic consequences of *Neisseria gonorrhoeae* pilin glycan variation. Microbes. Infect. **6**:693-701.

45. **Gibson, B. W., W. Melaugh, N. J. Phillips, M. A. Apicella, A. A. Campagnari, and J. M. Griffiss**. 1993. Investigation of the structural heterogeneity of

lipooligosaccharides from pathogenic *Haemophilus* and *Neisseria* species and of R-type lipopolysaccharides from *Salmonella typhimurium* by electrospray mass spectrometry. J. Bacteriol. **175**:2702-2712.

46. **Goodman, S. D. and J. J. Scocca**. 1988. Identification and arrangement of the DNA sequence recognized in specific transformation of *Neisseria gonorrhoeae*. Proc. Natl. Acad. Sci. U. S. A **85**:6982-6986.

47. **Goryshin, I. Y. and W. S. Reznikoff**. 1998. Tn5 in vitro transposition. J. Biol. Chem. **273**:7367-7374.

48. **Gotschlich, E. C.** 1994. Genetic locus for the biosynthesis of the variable portion of *Neisseria gonorrhoeae* lipooligosaccharide. J. Exp. Med. **180**:2181-2190.

49. **Gris, P., G. Vincke, J. P. Delmez, and J. P. Dierckx**. 1989. *Neisseria sicca* pneumonia and bronchiectasis. Eur. Respir. J. **2**:685-687.

50. **Gunn, J. S. and D. C. Stein**. 1996. Use of a non-selective transformation technique to construct a multiply restriction/modification-deficient mutant of *Neisseria gonorrhoeae*. Mol. Gen. Genet. **251**:509-517.

51. **Hammerschmidt, S., C. Birkholz, U. Zahringer, B. D. Robertson, P. J. van, O. Ebeling, and M. Frosch**. 1994. Contribution of genes from the capsule gene complex (cps) to lipooligosaccharide biosynthesis and serum resistance in *Neisseria meningitidis*. Mol. Microbiol. **11**:885-896.

52. **Harvey, H. A., M. P. Jennings, C. A. Campbell, R. Williams, and M. A. Apicella**. 2001. Receptor-mediated endocytosis of *Neisseria gonorrhoeae* into primary human urethral epithelial cells: the role of the asialoglycoprotein receptor. Mol. Microbiol. **42**:659-672.

53. **Heiddal, S., J. T. Sverrisson, F. E. Yngvason, N. Cariglia, and K. G. Kristinsson**. 1993. Native-valve endocarditis due to *Neisseria sicca*: case report and review. Clin. Infect. Dis. **16**:667-670.

54. **Herf, C., J. Nichols, S. Fruh, B. Holloway, and C. U. Anderson**. 1998. Meningococcal disease: recognition, treatment, and prevention. Nurse Pract. **23**:30, 33-30, 40.

55. **Hitchcock, P. J. and T. M. Brown**. 1983. Morphological heterogeneity among *Salmonella* lipopolysaccharide chemotypes in silver-stained polyacrylamide gels. J. Bacteriol. **154**:269-277.

56. **Hobbs, M. M., T. M. Alcorn, R. H. Davis, W. Fischer, J. C. Thomas, I. Martin, C. Ison, P. F. Sparling, and M. S. Cohen**. 1999. Molecular typing of *Neisseria gonorrhoeae* causing repeated infections: evolution of porin during passage within a community. J. Infect. Dis. **179**:371-381.

57. **Hobbs, M. M., A. Seiler, M. Achtman, and J. G. Cannon**. 1994. Microevolution within a clonal population of pathogenic bacteria: recombination, gene duplication and horizontal genetic exchange in the opa gene family of *Neisseria meningitidis*. Mol. Microbiol. **12**:171-180.

58. **Hoke, C. and N. A. Vedros**. 1982. Characterization of atypical aerobic gram-negative cocci isolated from humans. J. Clin. Microbiol. **15**:906-914.

59. **Hood, D. W., G. Randle, A. D. Cox, K. Makepeace, J. Li, E. K. Schweda, J. C. Richards, and E. R. Moxon**. 2004. Biosynthesis of cryptic lipopolysaccharide glycoforms in *Haemophilus influenzae* involves a mechanism similar to that required for O-antigen synthesis. J. Bacteriol. **186**:7429-7439.

60. **Inoue, H., H. Nojima, and H. Okayama**. 1990. High efficiency transformation of *Escherichia coli* with plasmids. Gene **96**:23-28.

61. **Jann, K., B. Jann, F. Orskov, I. Orskov, and O. Westphal**. 1965. [Immunochemical studies of K antigens from *Escherichia coli*. II. K antigen from E. coli 08:K42(A):H-]. Biochem. Z. **342**:1-22.

62. **Jennings, M. P., D. W. Hood, I. R. Peak, M. Virji, and E. R. Moxon**. 1995. Molecular analysis of a locus for the biosynthesis and phase-variable expression of the lacto-N-neotetraose terminal lipopolysaccharide structure in *Neisseria meningitidis*. Mol. Microbiol. **18**:729-740.

63. **John, C. M., J. M. Griffiss, M. A. Apicella, R. E. Mandrell, and B. W. Gibson**. 1991. The structural basis for pyocin resistance in *Neisseria gonorrhoeae* lipooligosaccharides. J. Biol. Chem. **266**:19303-19311.

64. **John, C. M., H. Schneider, and J. M. Griffiss**. 1999. *Neisseria gonorrhoeae* that infect men have lipooligosaccharides with terminal N-acetyllactosamine repeats. J. Biol. Chem. **274**:1017-1025.

65. **Jones, D. M., R. Borrow, A. J. Fox, S. Gray, K. A. Cartwright, and J. T. Poolman**. 1992. The lipooligosaccharide immunotype as a virulence determinant in *Neisseria meningitidis*. Microb. Pathog. **13**:219-224.

66. **Kahler, C. M., E. Blum, Y. K. Miller, D. Ryan, T. Popovic, and D. S. Stephens**. 2001. exl, an exchangeable genetic island in *Neisseria meningitidis*. Infect. Immun. **69**:1687-1696.

67. **Kahler, C. M., R. W. Carlson, M. M. Rahman, L. E. Martin, and D. S. Stephens**. 1996. Inner core biosynthesis of lipooligosaccharide (LOS) in *Neisseria meningitidis* serogroup B: identification and role in LOS assembly of the alpha1,2 N-acetylglucosamine transferase (RfaK). J. Bacteriol. **178**:1265-1273.

68. **Kahler, C. M., R. W. Carlson, M. M. Rahman, L. E. Martin, and D. S. Stephens**. 1996. Two glycosyltransferase genes, lgtF and rfaK, constitute the lipooligosaccharide ice (inner core extension) biosynthesis operon of *Neisseria meningitidis*. J. Bacteriol. **178**:6677-6684.

69. **Kahler, C. M., A. Datta, Y. L. Tzeng, R. W. Carlson, and D. S. Stephens**. 2005. Inner core assembly and structure of the lipooligosaccharide of *Neisseria meningitidis*: capacity of strain NMB to express all known immunotype epitopes. Glycobiology **15**:409-419.

70. **Kanipes, M. I., S. Lin, R. J. Cotter, and C. R. Raetz**. 2001. Ca2+-induced phosphoethanolamine transfer to the outer 3-deoxy-D-manno-octulosonic acid

moiety of *Escherichia coli* lipopolysaccharide. A novel membrane enzyme dependent upon phosphatidylethanolamine. J. Biol. Chem. **276**:1156-1163.

71. **Keenleyside, W. J. and C. Whitfield**. 1995. Lateral transfer of rfb genes: a mobilizable ColE1-type plasmid carries the rfbO:54 (O:54 antigen biosynthesis) gene cluster from *Salmonella enterica* serovar Borreze. J. Bacteriol. **177**:5247-5253.

72. **Kerwood, D. E., H. Schneider, and R. Yamasaki**. 1992. Structural analysis of lipooligosaccharide produced by *Neisseria gonorrhoeae*, strain MS11mk (variant A): a precursor for a gonococcal lipooligosaccharide associated with virulence. Biochemistry **31**:12760-12768.

73. **Kim, J. J., D. Zhou, R. E. Mandrell, and J. M. Griffiss**. 1992. Effect of exogenous sialylation of the lipooligosaccharide of *Neisseria gonorrhoeae* on opsonophagocytosis. Infect. Immun. **60**:4439-4442.

74. **Klee, S. R., X. Nassif, B. Kusecek, P. Merker, J. L. Beretti, M. Achtman, and C. R. Tinsley**. 2000. Molecular and biological analysis of eight genetic islands that distinguish *Neisseria meningitidis* from the closely related pathogen *Neisseria* gonorrhoeae. Infect. Immun. **68**:2082-2095.

75. **Kline, K. A. and H. S. Seifert**. 2005. Role of the Rep helicase gene in homologous recombination in *Neisseria gonorrhoeae*. J. Bacteriol. **187**:2903-2907.

76. **Kroll, J. S., K. E. Wilks, J. L. Farrant, and P. R. Langford**. 1998. Natural genetic exchange between *Haemophilus* and *Neisseria*: intergeneric transfer of chromosomal genes between major human pathogens. Proc. Natl. Acad. Sci. U. S. A **95**:12381-12385.

77. **Lansac, N., F. J. Picard, C. Menard, M. Boissinot, M. Ouellette, P. H. Roy, and M. G. Bergeron**. 2000. Novel genus-specific PCR-based assays for rapid identification of *Neisseria* species and *Neisseria meningitidis*. Eur. J. Clin. Microbiol. Infect. Dis. **19**:443-451.

78. **Lee, H., F. F. Hsu, J. Turk, and E. A. Groisman**. 2004. The PmrA-regulated pmrC gene mediates phosphoethanolamine modification of lipid A and polymyxin resistance in *Salmonella enterica*. J. Bacteriol. **186**:4124-4133.

79. **Li, M. S., J. L. Farrant, P. R. Langford, and J. S. Kroll**. 2003. Identification and characterization of genomic loci unique to the Brazilian purpuric fever clonal group of *H. influenzae* biogroup aegyptius: functionality explored using meningococcal homology. Mol. Microbiol. **47**:1101-1111.

80. **Linz, B., M. Schenker, P. Zhu, and M. Achtman**. 2000. Frequent interspecific genetic exchange between commensal *Neisseriae* and *Neisseria meningitidis*. Mol. Microbiol. **36**:1049-1058.

81. **Mackinnon, F. G., A. D. Cox, J. S. Plested, C. M. Tang, K. Makepeace, P. A. Coull, J. C. Wright, R. Chalmers, D. W. Hood, J. C. Richards, and E. R. Moxon**. 2002. Identification of a gene (lpt-3) required for the addition of

phosphoethanolamine to the lipopolysaccharide inner core of *Neisseria meningitidis* and its role in mediating susceptibility to bactericidal killing and opsonophagocytosis. Mol. Microbiol. **43**:931-943.

82. **Maier, B., M. Koomey, and M. P. Sheetz**. 2004. A force-dependent switch reverses type IV pilus retraction. Proc. Natl. Acad. Sci. U. S. A **101**:10961-10966.

83. **Mandrell, R. E., F. H. Azmi, and D. M. Granoff**. 1995. Complement-mediated bactericidal activity of human antibodies to poly alpha 2-->8 N-acetylneuraminic acid, the capsular polysaccharide of *Neisseria meningitidis* serogroup B. J. Infect. Dis. **172**:1279-1289.

84. **Mandrell, R. E., J. M. Griffiss, and B. A. Macher**. 1988. Lipooligosaccharides (LOS) of *Neisseria gonorrhoeae* and *Neisseria meningitidis* have components that are immunochemically similar to precursors of human blood group antigens. Carbohydrate sequence specificity of the mouse monoclonal antibodies that recognize crossreacting antigens on LOS and human erythrocytes. J. Exp. Med. **168**:107-126.

85. **Mandrell, R. E., A. J. Lesse, J. V. Sugai, M. Shero, J. M. Griffiss, J. A. Cole, N. J. Parsons, H. Smith, S. A. Morse, and M. A. Apicella**. 1990. In vitro and in vivo modification of *Neisseria gonorrhoeae* lipooligosaccharide epitope structure by sialylation. J. Exp. Med. **171**:1649-1664.

86. **McLeod, G. J., B. L. Brandt, N. B. Saunders, and W. Zollinger**. 2000. Structural relationships and sialylation among meningococcal L1, L8, and L3,7 lipooligosaccharide serotypes. J. Biol. Chem. **275**:9716-9724.

87. **Mehr, I. J. and H. S. Seifert**. 1998. Differential roles of homologous recombination pathways in *Neisseria gonorrhoeae* pilin antigenic variation, DNA transformation and DNA repair. Mol. Microbiol. **30**:697-710.

88. **Mirlashari, M. R. and T. Lyberg**. 2003. Expression and involvement of Toll-like receptors (TLR)2, TLR4, and CD14 in monocyte TNF-alpha production induced by lipopolysaccharides from *Neisseria meningitidis*. Med. Sci. Monit. **9**:BR316-BR324.

89. **Muhlecker, W., S. Gulati, D. P. McQuillen, S. Ram, P. A. Rice, and V. N. Reinhold**. 1999. An essential saccharide binding domain for the mAb 2C7 established for *Neisseria gonorrhoeae* LOS by ES-MS and MSn. Glycobiology **9**:157-171.

90. **Nassif, X., J. Lowy, P. Stenberg, P. O'Gaora, A. Ganji, and M. So**. 1993. Antigenic variation of pilin regulates adhesion of *Neisseria meningitidis* to human epithelial cells. Mol. Microbiol. **8**:719-725.

91. **Neu, A. M., B. Case, H. M. Lederman, and B. A. Fivush**. 1994. *Neisseria sicca* peritonitis in a patient maintained on chronic peritoneal dialysis. Pediatr. Nephrol. **8**:601-602.

92. **Orus, P. and M. Vinas**. 2001. Mechanisms other than penicillin-binding protein-2 alterations may contribute to moderate penicillin resistance in *Neisseria meningitidis*. Int. J. Antimicrob. Agents **18**:113-119.

93. **Perez-Losada, M., R. P. Viscidi, J. C. Demma, J. Zenilman, and K. A. Crandall**. 2005. Population Genetics of *Neisseria gonorrhoeae* in a High-Prevalence Community Using a Hypervariable Outer Membrane porB and 13 Slowly Evolving Housekeeping Genes. Mol. Biol. Evol. **22**:1887-1902.

94. **Perrin, A., S. Bonacorsi, E. Carbonnelle, D. Talibi, P. Dessen, X. Nassif, and C. Tinsley**. 2002. Comparative genomics identifies the genetic islands that distinguish *Neisseria meningitidis*, the agent of cerebrospinal meningitis, from other *Neisseria* species. Infect. Immun. **70**:7063-7072.

95. **Petricoin, E. F., III and D. C. Stein**. 1989. Molecular analysis of lipooligosaccharide biosynthesis in *Neisseria gonorrhoeae*. Infect. Immun. **57**:2847-2852.

96. **Phillips, N. J., M. A. Apicella, J. M. Griffiss, and B. W. Gibson**. 1992. Structural characterization of the cell surface lipooligosaccharides from a nontypable strain of *Haemophilus influenzae*. Biochemistry **31**:4515-4526.

97. **Plested, J. S., M. A. Gidney, P. A. Coull, H. G. Griffiths, M. A. Herbert, A. G. Bird, J. C. Richards, and E. R. Moxon**. 2000. Enzyme linked immunosorbent assay (ELISA) for the detection of serum antibodies to the inner core

lipopolysaccharide of *Neisseria meningitidis* group B. J. Immunol. Methods **237**:73-84.

98. **Plested, J. S., K. Makepeace, M. P. Jennings, M. A. Gidney, S. Lacelle, J. Brisson, A. D. Cox, A. Martin, A. G. Bird, C. M. Tang, F. M. Mackinnon, J. C. Richards, and E. R. Moxon**. 1999. Conservation and accessibility of an inner core lipopolysaccharide epitope of *Neisseria meningitidis*. Infect. Immun. **67**:5417-5426.

99. **Rahman, M., A. B. Jonsson, and T. Holme**. 1998. Monoclonal antibodies to the epitope alpha-Gal-(1-4)-beta-Gal-(1- of Moraxella catarrhalis LPS react with a similar epitope in type IV pili of *Neisseria meningitidis*. Microb. Pathog. **24**:299-308.

100. **Rahman, M. M., C. M. Kahler, D. S. Stephens, and R. W. Carlson**. 2001. The structure of the lipooligosaccharide (LOS) from the alpha-1,2-N-acetyl glucosamine transferase (rfaK(NMB)) mutant strain CMK1 of *Neisseria meningitidis*: implications for LOS inner core assembly and LOS-based vaccines. Glycobiology **11**:703-709.

101. **Rahman, M. M., D. S. Stephens, C. M. Kahler, J. Glushka, and R. W. Carlson**. 1998. The lipooligosaccharide (LOS) of *Neisseria meningitidis* serogroup B strain NMB contains L2, L3, and novel oligosaccharides, and lacks the lipid-A 4'-phosphate substituent. Carbohydr. Res. **307**:311-324.

102. **Ram, S., A. D. Cox, J. C. Wright, U. Vogel, S. Getzlaff, R. Boden, J. Li, J. S. Plested, S. Meri, S. Gulati, D. C. Stein, J. C. Richards, E. R. Moxon, and P. A. Rice**. 2003. Neisserial lipooligosaccharide is a target for complement component C4b. Inner core phosphoethanolamine residues define C4b linkage specificity. J. Biol. Chem. **278**:50853-50862.

103. **Raymond, C. K., E. H. Sims, A. Kas, D. H. Spencer, T. V. Kutyavin, R. G. Ivey, Y. Zhou, R. Kaul, J. B. Clendenning, and M. V. Olson**. 2002. Genetic variation at the O-antigen biosynthetic locus in *Pseudomonas aeruginosa*. J. Bacteriol. **184**:3614-3622.

104. **Rest, R. F. and J. V. Frangipane**. 1992. Growth of *Neisseria gonorrhoeae* in CMP-N-acetylneuraminic acid inhibits nonopsonic (opacity-associated outer membrane protein-mediated) interactions with human neutrophils. Infect. Immun. **60**:989-997.

105. **Reznikoff, W. S., I. Y. Goryshin, and J. J. Jendrisak**. 2004. Tn5 as a molecular genetics tool: In vitro transposition and the coupling of in vitro technologies with in vivo transposition. Methods Mol. Biol. **260**:83-96.

106. **Roberts, M. C., W. O. Chung, D. Roe, M. Xia, C. Marquez, G. Borthagaray, W. L. Whittington, and K. K. Holmes**. 1999. Erythromycin-resistant *Neisseria* gonorrhoeae and oral commensal *Neisseria* spp. carry known rRNA methylase genes. Antimicrob. Agents Chemother. **43**:1367-1372.

107. **Robertson, B. D., M. Frosch, and J. P. van Putten**. 1994. The identification of cryptic rhamnose biosynthesis genes in *Neisseria gonorrhoeae* and their relationship to lipopolysaccharide biosynthesis. J. Bacteriol. **176**:6915-6920.

108. **Sandlin, R. and D. C. Stein**. 1991. Structural heterogeneity of the lipopolysaccharides of the *Neisseriaceae*. FEMS Microbiol. Lett. **69**:69-72.

109. **Sandlin, R. C. and D. C. Stein**. 1994. Role of phosphoglucomutase in lipooligosaccharide biosynthesis in *Neisseria gonorrhoeae*. J. Bacteriol. **176**:2930-2937.

110. **Schilling, B., B. W. Gibson, M. Filiatrault, and A. A. Campagnari**. 2002. Characterization of lipooligosaccharides from *Haemophilus ducreyi* containing polylactosamine repeats. J. Am. Soc. Mass Spectrom. **13**:724-734.

111. **Schneider, H., A. S. Cross, R. A. Kuschner, D. N. Taylor, J. C. Sadoff, J. W. Boslego, and C. D. Deal**. 1995. Experimental human gonococcal urethritis: 250 *Neisseria gonorrhoeae* MS11mkC are infective. J. Infect. Dis. **172**:180-185.

112. **Schneider, H., J. M. Griffiss, J. W. Boslego, P. J. Hitchcock, K. M. Zahos, and M. A. Apicella**. 1991. Expression of paragloboside-like lipooligosaccharides may be a necessary component of gonococcal pathogenesis in men. J. Exp. Med. **174**:1601-1605.

113. **Schneider, H., J. M. Griffiss, R. E. Mandrell, and G. A. Jarvis**. 1985. Elaboration of a 3.6-kilodalton lipooligosaccharide, antibody against which is

absent from human sera, is associated with serum resistance of *Neisseria gonorrhoeae*. Infect. Immun. **50**:672-677.

114. **Schneider, H., K. A. Schmidt, D. R. Skillman, d. van, V, R. L. Warren, H. J. Wylie, J. C. Sadoff, C. D. Deal, and A. S. Cross**. 1996. Sialylation lessens the infectivity of *Neisseria gonorrhoeae* MS11mkC. J. Infect. Dis. **173**:1422-1427.

115. **Serino, L. and M. Virji**. 2002. Genetic and functional analysis of the phosphorylcholine moiety of commensal *Neisseria* lipopolysaccharide. Mol. Microbiol. **43**:437-448.

116. **Serino, L. and M. Virji**. 2000. Phosphorylcholine decoration of lipopolysaccharide differentiates commensal *Neisseriae* from pathogenic strains: identification of licA-type genes in commensal *Neisseriae*. Mol. Microbiol. **35**:1550-1559.

117. **Skaar, E. P., M. P. Lazio, and H. S. Seifert**. 2002. Roles of the recJ and recN genes in homologous recombination and DNA repair pathways of *Neisseria* gonorrhoeae. J. Bacteriol. **184**:919-927.

118. **Smith, J. M., N. H. Smith, M. O'Rourke, and B. G. Spratt**. 1993. How clonal are bacteria? Proc. Natl. Acad. Sci. U. S. A **90**:4384-4388.

119. **Smith, N. H., E. C. Holmes, G. M. Donovan, G. A. Carpenter, and B. G. Spratt**. 1999. Networks and groups within the genus *Neisseria*: analysis of argF, recA, rho, and 16S rRNA sequences from human *Neisseria* species. Mol. Biol. Evol. **16**:773-783.

120. **Smith, N. H., S. J. Maynard, and B. G. Spratt**. 1995. Sequence evolution of the porB gene of *Neisseria gonorrhoeae* and *Neisseria meningitidis*: evidence of positive Darwinian selection. Mol. Biol. Evol. **12**:363-370.

121. **Song, W., L. Ma, R. Chen, and D. C. Stein**. 2000. Role of lipooligosaccharide in Opa-independent invasion of *Neisseria gonorrhoeae* into human epithelial cells. J. Exp. Med. **191**:949-960.

122. **Stein, D. C.** 1991. Transformation of *Neisseria gonorrhoeae*: physical requirements of the transforming DNA. Can. J. Microbiol. **37**:345-349.

123. **Stevenson, G., B. Neal, D. Liu, M. Hobbs, N. H. Packer, M. Batley, J. W. Redmond, L. Lindquist, and P. Reeves**. 1994. Structure of the O antigen of Escherichia coli K-12 and the sequence of its rfb gene cluster. J. Bacteriol. **176**:4144-4156.

124. **Stojiljkovic, I., V. Hwa, J. Larson, L. Lin, M. So, and X. Nassif**. 1997. Cloning and characterization of the *Neisseria meningitidis* rfaC gene encoding alpha-1,5 heptosyltransferase I. FEMS Microbiol. Lett. **151**:41-49.

125. **Taha, M. K.** 2000. *Neisseria meningitidis* induces the expression of the TNF-alpha gene in endothelial cells. Cytokine **12**:21-25.

126. **Toleman, M., E. Aho, and M. Virji**. 2001. Expression of pathogen-like Opa adhesins in commensal *Neisseria*: genetic and functional analysis. Cell Microbiol. **3**:33-44.

127. **Tong, Y., D. Arking, S. Ye, B. Reinhold, V. Reinhold, and D. C. Stein**. 2002. *Neisseria gonorrhoeae* strain PID2 simultaneously expresses six chemically related lipooligosaccharide structures. Glycobiology **12**:523-533.

128. **Tong, Y., V. Reinhold, B. Reinhold, B. Brandt, and D. C. Stein**. 2001. Structural and immunochemical characterization of the lipooligosaccharides expressed by *Neisseria subflava* 44. J. Bacteriol. **183**:942-950.

129. **Troncoso, G., S. Sanchez, J. Kolberg, E. Rosenqvist, M. Veiga, C. M. Ferreiros, and M. Criado**. 2001. Analysis of the expression of the putatively virulence-associated neisserial protein RmpM (class 4) in commensal *Neisseria* and *Moraxella catarrhalis* strains. FEMS Microbiol. Lett. **199**:171-176.

130. **Tzeng, Y. L., A. Datta, K. Ambrose, M. Lo, J. K. Davies, R. W. Carlson, D. S. Stephens, and C. M. Kahler**. 2004. The MisR/MisS two-component regulatory system influences inner core structure and immunotype of lipooligosaccharide in *Neisseria meningitidis*. J. Biol. Chem. **279**:35053-35062.

131. **Valenzuela, G. A., T. D. Davis, E. Pizzani, and D. McGroarty**. 1992. Infective endocarditis due to *Neisseria sicca* and associated with intravenous drug abuse. South. Med. J. **85**:929.

132. **Verheul, A. F., H. Snippe, and J. T. Poolman**. 1993. Meningococcal lipopolysaccharides: virulence factor and potential vaccine component. Microbiol. Rev. **57**:34-49.

133. **Vogel, U., A. Weinberger, R. Frank, A. Muller, J. Kohl, J. P. Atkinson, and M. Frosch**. 1997. Complement factor C3 deposition and serum resistance in isogenic capsule and lipooligosaccharide sialic acid mutants of serogroup B *Neisseria meningitidis*. Infect. Immun. **65**:4022-4029.

134. **Wang, L., S. Huskic, A. Cisterne, D. Rothemund, and P. R. Reeves**. 2002. The O-antigen gene cluster of *Escherichia coli* O55:H7 and identification of a new UDP-GlcNAc C4 epimerase gene. J. Bacteriol. **184**:2620-2625.

135. **Westphal, O., K. Jann, and K. Himmelspach**. 1983. Chemistry and immunochemistry of bacterial lipopolysaccharides as cell wall antigens and endotoxins. Prog. Allergy **33**:9-39.

136. **White, L. A. and D. S. Kellogg, Jr.** 1965. *NEISSERIA GONORRHOEAE* IDENTIFICATION IN DIRECT SMEARS BY A FLUORESCENT ANTIBODY-COUNTERSTAIN METHOD. Appl. Microbiol. **13**:171-174.

137. **Wolff, K. and A. Stern**. 1995. Identification and characterization of specific sequences encoding pathogenicity associated proteins in the genome of commensal *Neisseria* species. FEMS Microbiol. Lett. **125**:255-263.

138. **Wolfgang, M., P. Lauer, H. S. Park, L. Brossay, J. Hebert, and M. Koomey**. 1998. PilT mutations lead to simultaneous defects in competence for natural transformation and twitching motility in piliated *Neisseria gonorrhoeae*. Mol. Microbiol. **29**:321-330.

139. **Wright, J. C., D. W. Hood, G. A. Randle, K. Makepeace, A. D. Cox, J. Li, R. Chalmers, J. C. Richards, and E. R. Moxon**. 2004. lpt6, a gene required for addition of phosphoethanolamine to inner-core lipopolysaccharide of *Neisseria meningitidis* and *Haemophilus influenzae*. J. Bacteriol. **186**:6970-6982.

140. **Yang, Q. L. and E. C. Gotschlich**. 1996. Variation of gonococcal lipooligosaccharide structure is due to alterations in poly-G tracts in lgt genes encoding glycosyl transferases. J. Exp. Med. **183**:323-327.

141. **Zhu, P., M. J. Klutch, M. C. Bash, R. S. Tsang, L. K. Ng, and C. M. Tsai**. 2002. Genetic diversity of three lgt loci for biosynthesis of lipooligosaccharide (LOS) in *Neisseria* species. Microbiology **148**:1833-1844.